Karl R. Popper

Post Scriptum a La lógica de la investigación científica. Vol. II

El universo abierto

Un argumento en favor del indeterminismo

Edición preparada por W.W. Bartley III

tecnos

SERIE DE FILOSOFÍA Y ENSAYO
Dirigida por Manuel Garrido

Alston, W. P.; Edwards, P.; Malcolm, N.; Nelson, J. O., y Prior, A.: *Los orígenes de la filosofía analítica. Moore, Russell, Wittgenstein.*
Austin, J. L.: *Sentido y percepción.*
Boden, M. A.: *Inteligencia artificial y hombre natural.*
Bottomore, T.; Harris, L.; Kiernan, V. G.; Miliband, R.; con la colaboración de Kolakowski, L.: *Diccionario del pensamiento marxista.*
Brown, H. I.: *La nueva filosofía de la ciencia.*
Bunge, M.: *El problema mente-cerebro.*
Couturat, Louis: *El álgebra de la lógica.*
Chrisholm, R. M.: *Teoría del conocimiento.*
Díaz, Elías: *Revisión de Unamuno.* Análisis crítico de su pensamiento político.
Doop, Joseph: *Nociones de lógica formal.*
Edelman, B.: *La práctica ideológica del Derecho.*
Fann, K. T.: *El concepto de filosofía en Wittgenstein.*
Ferrater Mora, J., y otros: *Filosofía y ciencia en el pensamiento español contemporáneo (1960-1970).*
Feyerabend, P.: *Tratado contra el método.* Esquema de una teoría anarquista del conocimiento.
García Suárez, A.: *La lógica de la experiencia.*
Garrido, M.: *Lógica simbólica.*
Gómez García, P.: *La antropología estructural de Claude Lévi-Strauss.*
Gurméndez, Carlos: *Ser para no ser.* Ensayo de una dialéctica subjetiva.
Hierro, José S.-P.: *Problemas del análisis del lenguaje moral.*
Hintikka, J.: *Lógica, juegos de lenguaje e información.* Temas kantianos de filosofía de la lógica.
Hintikka, J.: *Saber y creer.* Una introducción a la lógica de las dos nociones.
Kuhn, T. S.: *Segundos pensamientos sobre paradigmas.*
Lakatos, Imre, y otros: *Historia de la ciencia y sus reconstrucciones racionales* (2.ª ed.).
Lindsay, P. H., y Norman, D. A.: *Introducción a la psicología cognitiva* (2.ª ed.).
Lorenzen, Paul: *Metamatemática.*
Lorenzo, J. de: *El método axiomático y sus creencias.*
Lorenzo, J. de: *La filosofía de la matemática de Jules Henri Poincaré.*
Martín Santos, Luis, y otros: *Ensayos de filosofía de la ciencia.* Simposio de Burgos. En torno a la obra de Karl R. Popper.
Mates, Benson: *Lógica matemática elemental.*
París, Carlos: *Hombre y naturaleza.*
Popper, Karl R.: *Búsqueda sin término.* Una autobiografía intelectual.
Popper, Karl R.: *Realismo y el objetivo de la ciencia.* Post Scriptum a la lógica de la investigación científica, vol. I.
Popper, Karl R.: *El universo abierto.* Post Scriptum a la lógica de la investigación científica, vol. II.
Popper, Karl R.: *Teoría científica y el cisma en física.* Post Scriptum a la lógica de la investigación científica, vol. III.
Prior, A. N.: *Historia de la lógica.*
Quine, W. V.: *La relatividad ontológica y otros ensayos.*
Quintanilla, Miguel A.: *Idealismo y filosofía de la ciencia.* Introducción a la epistemología de Karl R. Popper.
Rama, Carlos M.: *Teoría de la historia.* Introducción a los estudios históricos (3.ª ed.).
Rescher, Nicolas: *La primacía en la práctica.*
Robinet, André: *Mitología, filosofía y cibernética.*
Rodríguez Paniagua, José M.ª: *¿Derecho natural o axiología jurídica?*
Rodríguez Paniagua, José M.ª: *Marx y el problema de la ideología.*
Sahakian, W. S.: *Historia y sistemas de la psicología.*
Smart, J. J. C., y Williams, B.: *Utilitarismo: pro y contra.*
Sotelo, Ignacio: *Sartre y la razón dialéctica.*
Strawson, P. F.: *Ensayos lógico-lingüísticos.*
Vargas Machuca, R.: *El poder moral de la razón.* La filosofía de Gramsci.
Veldman, D. J.: *Programación de computadoras en ciencias de la conducta.*
Wellman, C.: *Morales y éticas.*

EL UNIVERSO ABIERTO

KARL R. POPPER

Post Scriptum a La lógica de la investigación científica

Volumen II

EL UNIVERSO ABIERTO

UN ARGUMENTO EN FAVOR DEL INDETERMINISMO

Edición preparada por
W. W. Bartley III

tecnos

Título original:
The Open Universe. An Argument for Indeterminism.
From the Postscript to the Logic of Scientific Discovery
Edited by W. W. Bartley III

Traducción:
MARTA SANSIGRE VIDAL

Diseño de cubierta:
J. M. DOMINGUEZ y J. SANCHEZ CUENCA

Impresión de cubierta:
GRAFICAS MOLINA

1.ª edición, 1984
1.ª reimp., 1986

Reservados todos los derechos. Ni la totalidad ni parte de este libro puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico, incluyendo fotocopia, grabación magnética o cualquier almacenamiento de información y sistema de recuperación, sin permiso escrito de Editorial Tecnos, S. A.

© Karl R. Popper, 1956, 1982
© EDITORIAL TECNOS, S. A., 1986
O'Donnell, 27 - 28009 Madrid
ISBN: 84-309-1059-X
Depósito Legal: M-8908-1986

Printed in Spain. Impreso en España por Gama, S. A.
Tracia, 17. Madrid

A Ernst Gombrich

Digitized by the Internet Archive
in 2022 with funding from
Kahle/Austin Foundation

https://archive.org/details/postscriptumlalo0002popp

INDICE

PRINCIPALES PUBLICACIONES Y ABREVIATURAS DE TÍTULOS ..... *Pág.* 11

PROLOGO, por W. W. Bartley III 13

AGRADECIMIENTOS 19

PREFACIO DE 1982 21

CAPÍTULO I: CLASES DE DETERMINISMO 25
1. (*62) Determinismo: religioso, «científico» y metafísico 28
2. (*63) Preguntas de-por-qué. Causalidad y determinismo «científico» 32
3. (*64) El principio de poder dar razón 35
4. (*65) El estudio de la conducta y el principio de poder dar razón 38
5. (*66) Temperaturas críticas y el principio de todo-o-nada ... 40
6. (*67) Relojes y nubes 41
7. (*68) Argumentos desde la psicología en favor del determinismo 43
8. (*69) La visión determinista del mundo 48
9. (*70) La carga de la prueba 50

CAPÍTULO II: DETERMINISMO «CIENTIFICO» 52
10. (*71) El determinismo *prima facie* de la física clásica. El demonio de Laplace 52
11. (*72) La idea de determinismo «científico»: predictibilidad desde dentro 55
12. (*73) Dos definiciones de determinismo «científico» 59
13. (*74) ¿Se desprende el determinismo «científico» de una teoría determinista *prima facie?* 60
14. (*75) Un resultado de Hadamard 61

CAPÍTULO III: EL ARGUMENTO EN FAVOR DEL INDETERMINISMO 64
15. (*76) Por qué soy indeterminista: las teorías como redes 64
16. (*77) Comparación con la perspectiva de Kant 70

Nota sobre la numeración de las secciones. Las secciones del *Post Scriptum* están numeradas consecutiva e independientemente dentro de cada uno de los tres tomos. Los números originales de las secciones indican el orden de éstas dentro de la totalidad del *Post Scriptum* y van aquí entre paréntesis, precedidas de un asterisco. *Ed.*

17. (*78) ¿Es la física clásica capaz de dar razón? 72
18. (*79) El pasado y el futuro 78
19. (*80) El veredicto de la relatividad especial 80
20. (*81) Predicción histórica y el aumento del conocimiento. 85
21. (*82) Predicción del aumento del conocimiento teórico 87
22. (*83) La imposibilidad de la autopredicción 91
23. (*84) La refutación del determinismo «científico» 100
24. (*85) Un argumento de San Agustín, Descartes y Haldane. 104

CAPÍTULO IV: CUESTIONES METAFISICAS 109
25. (*86) Las doctrinas metafísicas de determinismo e indeterminismo 109
26. (*87) Por qué rechazo el determinismo metafísico. Una conversación con Parménides 111
27. (*88) El beneficio para la ciencia: una teoría de las propensiones 114
28. (*89) Teorías deterministas *prima facie* y teorías probabilistas 117
29. (*90) La cuchilla de Landé 121
30. (*91) La cuchilla de Landé y la interpretación propensivista. 126
31. (*92) Conclusión 128

ADDENDA

EL INDETERMINISMO NO BASTA: UN EPILOGO 134
Los Mundos 1, 2 y 3 135
La realidad de los tres Mundos 137
La realidad de los Mundos 1 y 2 137
La realidad y parcial autonomía del Mundo 3 139
La situación humana y el mundo natural 143
Determinismo e indeterminismo en física 144
El indeterminismo no basta 147
Determinismo y naturalismo 148
El universo abierto 149

REDUCCION CIENTIFICA Y LA INCOMPLETUD ESENCIAL DE TODA CIENCIA 152

MAS OBSERVACIONES SOBRE LA REDUCCION, 1981 183

INDICE DE NOMBRES 199

INDICE DE MATERIAS 201

PRINCIPALES PUBLICACIONES Y ABREVIATURAS DE TITULOS

En las *Notas* se han utilizado las siguientes abreviaturas para referirse a las principales publicaciones del autor:

L. d. F.:	*Logik der Forschung,* Julius Spriner Verlag, Viena, 1934.
O. S.:	*The Open Society and Its Enemies,* Princeton University Press, 1950. Traducida como *La sociedad abierta y sus enemigos.* Editorial Paidós, Buenos Aires, 1957.
The Poverty:	*The Poverty of Historicism* (1944-5), 1957, 1960. Traducida como *La miseria del historicismo.* Taurus Ediciones, Madrid, 1961.
L. Sc. D.:	*The Logic of Scientific Discovery,* Hutchinson & Co., Londres; Basic Books, Nueva York, 1959. Traducida como *La lógica de la investigación científica.* Editorial Tecnos, Madrid, 1962.
C. & R.:	*Conjectures and Refutations: The Growth of Scientific Knowledge,* Routledge and Kegan Paul, Londres; Basic Books, Nueva York, 1963. Traducida como *El desarrollo del conocimiento científico: Conjeturas y refutaciones.* Editorial Paidós, Buenos Aires, 1967.
Obj. Kn.:	*Objective Knowledge: An Evolutionary Approach,* Clarendon Press, Oxford, 1972. Traducida como *Conocimiento objetivo: Un enfoque evolucionista.* Editorial Tecnos, 1974.
U. Q.:	*Unended Quest: An Intellectual Autobiography,* Fontana/Collins, Londres, 1976. Traducida como *Búsqueda sin término: Una autobiografía intelectual.* Editorial Tecnos, Madrid, 1977.
S. & I. B.:	*The Self and Its Brain,* en colaboración con Sir John Eccles, 1977. Traducida como *El yo y su cerebro.* Editorial Labor, Barcelona, 1980.
Q. T. & S. Ph.:	*Quantum Theory and the Schism in Phisics,* Vol. I of the *Postscript to The Logic of Scientific Discovery.* [*Teoría cuántica y el cisma en física* (traducción en preparación).]
R. & A. of Sc.:	*Realism and the Aim of Science.* Vol. III of the *Postscript to The Logic of Scientific Discovery.* [*Realismo y el objetivo de la ciencia* (traducción en preparación).]

PROLOGO

Este libro, El universo abierto: Un argumento en favor del indeterminismo, *es el segundo tomo, y contiene el núcleo del argumento, del tan esperado* Post Scriptum a La lógica de la investigación científica *de Sir Karl Popper. Aunque se escribió hace unos veinticinco años, no se había publicado todavía. No obstante, en él se contiene el tratamiento más importante y mejor sustentado de los problemas del determinismo y del indeterminismo de los que yo tengo noticia.*

El Post Scriptum a La lógica de la investigación científica *se escribió principalmente durante los años de 1951 a 1956, en el momento en que el primer libro publicado de Popper,* Logik der Forshung *(1934), se estaba traduciendo al inglés con el título de* The Logic of Scientific Discovery.

Los distintos volúmenes del Post Scriptum *formaban parte, originariamente, de una serie de Apéndices a* La lógica de la investigación científica, *con los que Popper se proponía corregir, extender y desarrollar las ideas de su primer libro. Algunos de esos Apéndices se incluyeron finalmente en* La lógica de la investigación científica, *cuando se publicó en 1959. Pero un grupo de Apéndices adquirió vida propia y se fue convirtiendo, gradualmente, en un trabajo bien integrado y distinto, que excedía en longitud al original* Logik der Forschung. *Se decidió publicar esta nueva obra —titulada* Postscript: After Twenty Years *(Post Scriptum: Veinte años después)— como corolario o volumen compañero de* La lógica de la investigación científica, *y, en consecuencia, se imprimió, en galeradas, en 1956-57.*

Pocos meses antes de la fecha prevista para la publicación, sin embargo, el proyecto se paró en seco. En Búsqueda sin término, *su autobiografía intelectual, Sir Karl dice de esas galeradas: «La corrección de pruebas se convirtió en una pesadilla... Entonces tuve que someterme a una operación en ambos ojos. Después de ella no pude entregarme de nuevo a la corrección de pruebas hasta pasado algún tiempo y, como resultado, el* Postscript *sigue aún sin publicar.»*

Recuerdo claramente aquellos momentos. Fui a Viena a visitar a Popper en el hospital poco después de su operación por varios desprendimientos en las dos retinas, y trabajamos sobre el Post Scriptum *mientras él se iba recuperando. Durante mucho tiempo apenas veía y temíamos que se quedara ciego.*

Cuando recobró la vista, trabajó mucho en el Post Scriptum: *añadió varias secciones e hizo miles de correcciones a las galeradas. Pero la presión del resto de su trabajo se hacía excesiva y después de 1962 apenas se añadió nada al texto. Durante la década siguiente, enormemente productiva, Popper publicó* El desarrollo del conocimiento científico: Conjeturas y refutaciones *(1963), y completó y publicó tres libros nuevos:* Conocimiento objetivo: Un enfoque evolucionista *(1972),* Búsqueda sin término *(1974 y 1976) y, con Sir John Eccles,* El yo y su cerebro *(1977), así como numerosos artículos. Esos fueron los años, y las obras, en los que su ahora famosa teoría de la mente objetiva (y de los «Mundos 1, 2 y 3») se desarrolló y en los que su enfoque se extendió a las ciencias biológicas.*

Mientras tanto, el Post Scriptum, *que representaba la culminación del trabajo de Sir Karl en la filosofía de la física, seguía sin publicarse. Pero no sin leerse: la mayor parte de los alumnos y colegas más allegados de Popper han estudiado esta obra a lo largo de los años y varios han poseído copias de las galeradas. Para los que, como yo, han conocido este libro y han sido profundamente influidos por él es una gran satisfacción verlo finalmente completo y compartirlo con el gran público.*

El texto que se ha preparado ahora para su publicación es, esencialmente, el que existía en 1962. Excepto en algunos lugares, que se han señalado, no se han hecho alteraciones importantes. Pensamos que éste era el enfoque apropiado para una obra que ha adquirido, a través de su influencia en los alumnos y los colegas de Popper, un carácter histórico, al haber transcurrido unos veinticinco años desde que se redactó, y cuarenta y cinco desde que se escribió la Logik der Forschung *original. Es obvio que muchos puntos se hubieran expresado hoy de otra forma. Pero una revisión completa por parte del autor habría retrasado indefinidamente la publicación del libro.*

La preparación para la publicación ha consistido en reunir las diferentes versiones de algunas partes del texto que se habían ido acumulando a lo largo de los años; en revisar el libro, y en añadir notas bibliográficas y de otra índole para ayudar al lector. Las

pocas nuevas adiciones del propio Popper aparecen marcadas claramente: están entre corchetes y señaladas con un asterisco. Mis breves notas, bibliográficas y editoriales aparecen también entre corchetes, seguidas de la abreviatura «Ed.». Aquí he seguido, en general, la práctica establecida por Troels Eggers Hansen, editor de Die beiden Grundprobleme der Erkenntnistheorie *de Popper (escrita en 1930-32 y publicada en 1979). Popper ha podido comprobar el trabajo editorial en una serie de reuniones que hemos tenido en distintos lugares en los dos años últimos: en Heidelberg, Guelph, Toronto, Washington D. C., Schloss Kronberg y en su casa de Buckinghamshire. Popper ha añadido nuevos prefacios a todos los volúmenes, y un nuevo epílogo al segundo.*

Se ha hecho un cambio importante en la presentación, por sugerencia mía. Publicar esta extensa obra en un solo tomo hubiera sido posible, pero habría resultado un libro pesado y difícil de manejar, inasequible para muchos estudiantes de filosofía. Algunas partes del Post Scriptum *—incluido* El universo abierto: Un argumento en favor del indeterminismo— *tienen un amplio interés, no sólo para filósofos y estudiantes de filosofía, sino también para un público más extenso.*

Estas partes son también, en su totalidad, independientes unas de otras. Esto me llevó a sugerir que se publicase la obra en tres tomos independientes del mismo formato, que en conjunto integrasen el Post Scriptum. *Después de algunas dudas, Sir Karl aceptó la propuesta y también los títulos que yo le sugerí para los tres tomos.*

El Post Scriptum *se publica, pues, como sigue:*

Realismo y el objetivo de la ciencia *(tomo I).*

El Universo abierto: Un argumento en favor del indeterminismo *(tomo II).*

Teoría cuántica y el cisma en física *(tomo III).*

Aunque los tres volúmenes del Post Scriptum *pueden leerse por separado, el lector debe saber que forman un argumento unitario. Cada volumen del* Post Scriptum *ataca uno u otro de los enfoques subjetivistas o idealistas del conocimiento; cada uno estructura un componente o más para un enfoque objetivo y realista del conocimiento.*

Así, en el volumen titulado Realismo y el objetivo de la ciencia, *Popper acosa al «inductivismo», al que considera como principal fuente del subjetivismo e idealismo, en cuatro planos: lógico, metodológico, epistemológico y metafísico. Popper desarrolla su*

teoría de la falsabilidad y traza sus efectos, deslindando las opiniones científicas de las no científicas y de las pseudo-científicas. Y presenta su teoría sobre la corroboración como una forma de expresar preferencia racional por una teoría frente a otra sin recurrir a las «certezas» subjetivas o a la «justificación» de las filosofías convencionales. En el primer tomo, Popper trata también su relación con los personajes históricos de la filosofía, tales como Berkeley, Hume, Kant, Mach y Russell, que han hecho contribuciones importantes a la tradición subjetivista; y da respuestas detalladas a críticos filósofos y científicos contemporáneos. Popper ataca, asimismo, la interpretación subjetiva del cálculo de probabilidades, una interpretación que tiene sus raíces en la creencia de que la probabilidad mide un estado subjetivo de conocimiento insuficiente. En La lógica de la investigación científica, *Popper defendía una interpretación objetiva del cálculo de probabilidades, para lo que utilizaba la interpretación frecuencial. Ahora Popper critica también la interpretación frecuencial. En su lugar, presenta con todo detalle su propia interpretación propensivista, una interpretación que, durante los últimos veinte años, ha encontrado muchos defensores. Estas ideas y estos argumentos se aplican y se desarrollan en los volúmenes restantes.*

En El Universo abierto: Un argumento en favor del indeterminismo, *Popper presenta la crítica tanto de la forma «científica» como de la forma metafísica del determinismo y arguye que la física clásica no presupone ni implica determinismo más de lo que lo implica la física cuántica. Y, sin embargo, encuentra que el determinismo metafísico continúa subyacente en la obra de muchos teóricos cuánticos contemporáneos, incluidos los oponentes del determinismo. Popper encuentra el origen del papel que siguen jugando dentro de la física las interpretaciones subjetivas de la probabilidad en esas presuposiciones metafísicas deterministas.*

Hay una profunda conexión entre los argumentos de los tomos primero y segundo, en la preocupación que muestran por la libertad, la creatividad y la racionalidad del hombre.

El primer tomo, en su consideración de la justificación y de la racionalidad, rebate una pretensión subjetivista y escéptica sobre los límites de la crítica, y, con ello, los límites de la racionalidad. Si existiera ese límite, sería inútil el debate serio, y su apariencia sería ilusoria.

El segundo tomo, en su tratamiento del determinismo, defiende la idea de que nuestra racionalidad está limitada con respecto

a la predicción del aumento futuro del conocimiento humano. Si tal límite no *existiera, entonces sería inútil el debate serio, y su apariencia sería ilusoria.*

Popper mantiene, pues, que la razón humana no tiene límites con respecto a su capacidad de crítica y, sin embargo, los tiene respecto a su capacidad de predicción; y demuestra que la falta de limitación, en su caso, y la limitación, en otro, son absolutamente necesarias para que la racionalidad humana exista.

En el tomo III, Teoría cuántica y el cisma en física, *Popper reexamina y rebate una serie de argumentos y «paradojas» extensamente utilizados para defender un punto de vista idealista. Popper conjetura que los problemas de interpretación de la mecánica cuántica pueden reducirse a problemas de interpretación del cálculo de probabilidades, y a partir de ahí desarrollar con más amplitud su propia interpretación propensivista de la probabilidad. Y a continuación expone su extensa crítica de algunas de las interpretaciones dominantes de la teoría cuántica, tratando de resolver sus conocidas paradojas y de exorcizar al «observador» de la física cuántica.*

En su concluyente «Epílogo metafísico», Popper teje los temas de todo el Post Scriptum *en un estudio histórico y programático del papel de los programas de investigación metafísica o de las interpretaciones en la historia de la física.*

El editor quiere expresar su gratitud a The American Council of Learned Societies y a The American Philosophical Society por el generoso apoyo prestado a su trabajo editorial en estos volúmenes; y a los profesores Donald T. Campbell y F. A. von Hayek por sus consejos y su apoyo. También desea agradecer a su secretaria, Nancy Artis Sadoyama, su dedicación e incansable ayuda.

W. W. BARTLEY III

AGRADECIMIENTOS

Quiero aprovechar esta oportunidad para agradecer a mi colega W. N. Watkins el ánimo que me ha dado con su constante interés. Ha leído este volumen en manuscrito y en galeradas y ha hecho sugerencias muy útiles para mejorarlo. Fue por sugerencia suya por que decidí publicar el *Post Scriptum* como una obra aparte, en lugar de como había pensado originariamente, como una serie de Apéndices a *La lógica de la investigación científica.* Pero aún más importante para la terminación de la obra fue el interés que mostró por las ideas que contiene.

Quiero también dar las gracias a los cotraductores de *The Logic of Scientific Discovery,* doctor Julius Freed y Lan Freed, que leyeron la mayor parte de este volumen en galeradas e hicieron numerosas sugerencias para mejorar el estilo. [Los dos murieron muchos años antes de que se publicara el libro. *Ed.*]

Joseph Agassi fue, durante el tiempo en que se escribió este libro, primero estudiante de investigación y luego mi ayudante de investigación. Con él discutí casi cada sección con todo detalle, y con frecuencia, como consecuencia de sus consejos, desarrollé uno o dos enunciados hasta formar una nueva sección completa o, en su caso, una nueva parte [la parte 2 de *Realismo y el objetivo de la ciencia*]. Su cooperación fue para mí de enorme valor.

Quiero dar las gracias también a la London School of Economics and Political Science, que me hizo posible la provechosa ayuda del doctor Agassi, y al Center for Advanced Study in the Behavioral Sciences (Ford Foundation), en Stanford, California, por darme la oportunidad de trabajar ininterrumpidamente, desde octubre de 1956 a julio de 1957, en las galeradas de este libro, y por hacer posible al doctor Agassi que me ayudara durante ese tiempo.

PENN, BUCKINGHAMSHIRE, 1959

El profesor W. W. Bartley III fue alumno mío y, más tarde, colega en la London School of Economics, de 1958 a 1963, y colaboró estrechamente conmigo en 1960-62 en este libro. En 1978 tuvo la amabilidad de consentir en actuar como editor del *Post Scriptum.* Le estoy muy agradecido por su asistencia y por emprender tan ardua tarea. Le debo más de lo que puedo expresar.

Es también un placer dar las gracias a varias personas más, que, en los años pasados, trabajaron conmigo en este *Post Scriptum,* en particular, Alan E. Musgrave, David Miller, Arne F. Petersen, Tom Settle y Jeremy Shearmur. De ellos, David Miller y Arne F. Petersen deben ser mencionados especialmente a causa de la enorme cantidad de trabajo que ambos realizaron en varios períodos antes de 1970.

La London School ha continuado ayudándome durante todos estos años, designando un ayudante de investigación. Durante los trece años transcurridos desde mi jubilación, en 1969, lo ha hecho con ayuda de una beca de la Nuffield Foundation, a la que quiero expresar mi agradecimiento. Los principales responsables de esta disposición fueron: mi amigo y sucesor el profesor John Watkins; el difunto sir Walter Adams, director de la London School, y el director actual, profesor Ralf Dahrendorf, por cuya entrañable amistad e interés por mi trabajo: estoy profundamente agradecido.

Si el *Post Scriptum* se hubiera publicado en la década de 1950, se lo habría dedicado a Bertrand Russell: el profesor Bartley me ha dicho que hay una carta que lo atestigua en los Archivos de Russell en McMaster University.

Debo mencionar, finalmente, que el *Post Scriptum* (junto con la traducción inglesa de *La lógica de la investigación científica)* me parecía casi terminado en 1954. Fue entonces cuando escogí su título original «Postscript: After Twenty Years» [Post Scriptum: Veinte años después] aludiendo a la publicación de *Logik der Forschung* en 1934.

PENN, BUCKINGHAMSHIRE, 1982

PREFACIO DE 1982

Como explica el profesor Bartley en su Prólogo, este libro se escribió antes de 1956, como parte de mi *Post Scriptum a La lógica de la investigación científica.* En él no me proponía tratar sobre la libertad ni la libre voluntad humanas, a pesar de que ésos eran realmente los problemas de fondo.

Había varias razones por las que no quería tratar explícitamente esas cuestiones cuando escribí este libro. La razón principal era que quería que el *Post Scriptum,* igual que *La lógica de la investigación científica,* libro del que es continuación, tratase de las ciencias físicas, sus métodos y algunas de sus implicaciones, la cosmología física y el papel de la teoría del conocimiento en las ciencias físicas.

Una segunda razón es que yo tenía la impresión de que algunos de los problemas en torno al de la libertad humana han resultado en cierto modo embrollados por lo que los filósofos han escrito sobre ellos. Este embrollo empieza por lo menos con Hume, uno de los más grandes filósofos de todos los tiempos. Yo creo que todo está relacionado con el hecho de que el sentido común está enredado en esos problemas. Porque 1) el sentido común se inclina, por un lado, a afirmar que *todo* suceso es causado por algunos de los sucesos que lo preceden de modo que todo suceso puede ser explicado o predicho si conocemos todos los sucesos precedentes significativos con suficiente detalle. Por otro lado, 2) el sentido común atribuye a las personas maduras y cuerdas, al menos en muchas situaciones, la capacidad de escoger libremente entre posibilidades alternativas de acción y, por tanto, las atribuye responsabilidad de dicha acción.

Ahora bien, 1) y 2) parecen chocar, y el problema es si el choque es real o aparente. Hume y muchos deterministas han intentado, con ingenio y sutileza (en mi opinión, con demasiada sutileza), demostrar que 1) y 2) no chocan, en realidad, de forma que uno puede ser determinista y al mismo tiempo hablar de libertad de acción. Los argumentos en los que se basa esta posición son, sin embargo, principalmente verbales. Dependen del análisis

verbal del significado de palabras, tales como «libre», «voluntad» y «acción»; y del análisis de preguntas tales como «¿Podría yo haber hecho otra cosa que la que hice?». Estos análisis verbales son bastante fútiles y han llevado a la filosofía moderna a un estado de confusión. Pero hay otro planteamiento.

Mi propio punto de partida, y mi planteamiento de esas cuestiones, no proviene de Hume (que ha dominado el debate británico de ellas), sino de Laplace. En su *Essai philosophique sur les probabilités*, 1819, publicado unos cincuenta años después del *Treatise* de Hume, Laplace escribió:

> Debiéramos... considerar el presente estado del universo como el efecto de su estado anterior y la causa del que seguiría. Supongamos... una inteligencia que pudiera conocer todas las fuerzas que animan la naturaleza, y los estados, en un instante, de todos los objetos que la componen; ...para [esta inteligencia] nada podría ser incierto; y el futuro, como el pasado, sería presente a sus ojos [1].

A esta posición, que es (como explico en este libro) mucho más fuerte que el sentido común, la llamo determinismo «científico». Es una posición con la que la mayoría de los físicos, aunque no todos (estoy pensando, por ejemplo, en Charles Sanders Peirce), hubieran estado de acuerdo, por lo menos antes de 1927. Este determinismo laplaciano afirma que el estado del universo en un momento dado, futuro o pasado, está completamente determinado si su estado, su situación, es dado en algún momento, por ejemplo, el momento presente. Yo considero este determinismo laplaciano —aunque parezca confirmado por las teorías deterministas *prima facie* de la física y por sus maravillosos éxitos— como el obstáculo más sólido y serio en el camino de la explicación y la defensa de la libertad, la creatividad y la responsabilidad humanas. Ciertamente, el claro y firme enunciado determinista de Laplace, que va más allá del sentido común y que está también profundamente entrelazado con la historia de la ciencia occidental, es muy superior a la vaga formulación dada más arriba como 1).

El determinismo laplaciano puede, sin embargo, ser rebatido, y yo lo rebato en este libro. Yo, por mi parte, no soy determinista. Y me he propuesto en este libro abrir espacio para el indetermi-

[1] Traducido al inglés como *A Philosophical Essay on Probabilities*, 1951, páginas 4-5.

nismo dentro de la teoría física y dentro de la cosmología. Así, mantengo que el determinismo laplaciano es insostenible y, además, que no lo requieren ni la física «clásica» ni la contemporánea. Este es un cometido serio, que no tiene nada que ver con subterfugios verbales [2]. Mi argumentación, pues, será más en un plano cosmológico: hablaré del carácter de nuestro mundo en vez de hablar del significado de las palabras.

Pero, para evitar malentendidos, quiero dejar bien claro que cuando hablo del determinismo «científico» (siempre entre comillas) estoy pensando en una doctrina *supuestamente* «científica», en un determinismo *supuestamente* «científico». Es una doctrina que debe su popularidad y su influencia, incluso entre los grandes científicos, a su carácter aparentemente científico, al hecho de que se la considere por muchos como parte de la ciencia, especialmente de las teorías de la gravitación de Newton y de Einstein y de la teoría de Maxwell sobre el campo electromagnético.

A lo largo de mi argumentación desarrollaré, en oposición al determinismo, una oposición que trata también de ir más allá de la formulación citada más arriba como 2).

En cualquier caso, quiero afirmar aquí, claramente, algo que es manifiesto, tanto en *La sociedad abierta y sus enemigos* como en *La miseria del historicismo:* que estoy profundamente interesado en la defensa de la libertad humana, de la creatividad humana y de lo que se ha llamado tradicionalmente el libre albedrío (o libre voluntad), aunque creo que preguntas tales como «¿Qué es la libertad?» o «¿Qué significa 'libre'?» y «¿Qué es voluntad?» y otras similares pueden conducir al embrollo de la filosofía del lenguaje. Este libro es, pues, una especie de prolegómeno a la cuestión de la libertad y la creatividad humanas, y les abre espacio física y cosmológicamente de una manera que no depende de análisis verbales.

Respecto al comienzo del primer capítulo del libro, conviene decir unas palabras sobre mi opinión acerca de la intuición.

---

[2] Me permito recordar a mis lectores que evitar cuestiones verbales ha sido uno de los principios que me han guiado desde el comienzo de mi carrera. Véase el Prefacio a la 1.ª ed., 1934, de *La lógica de la investigación científica* [*L. d. F.*], especialmente las citas de Schlick y Kant. Véase también *La sociedad abierta y sus enemigos* [*O. S.*, 1945], capítulo 11, y *Búsqueda sin término* [*U. Q.*, 1976], sección 7. [Las siglas y las fechas entre corchetes corresponden a los títulos originales y a sus fechas de publicación. Los datos completos de publicación en el idioma original y en castellano se dan al final de esta obra. *N. de la T.*]

Considero que la intuición y la imaginación son sumamente importantes: las necesitamos para inventar una teoría. Pero la intuición, como puede persuadirnos y convencernos de la verdad de lo que hemos intuido, puede engañarnos de mala manera: es una ayuda inapreciable, pero también peligrosa, porque suele hacernos perder el sentido crítico. Debemos acercarnos a ella con respeto, con gratitud y con un esfuerzo por ser sus críticos severos.

CAPITULO PRIMERO

CLASES DE DETERMINISMO

Quiero exponer aquí las razones que tengo para ser indeterminista [1]. No incluiré entre ellas la idea del libre albedrío: como argumento racional en favor del indeterminismo es inútil. Un hombre puede creer que está actuando deliberadamente y por su propia y libre elección, cuando en realidad está actuando bajo la influencia de una sugerencia, de una compulsión o de las drogas. Pero, una vez que hayamos conseguido rechazar la idea de determinismo con argumentos que no apelen a nuestras intuiciones respecto al libre albedrío, quizá sea posible restablecer, sólo hasta cierto punto, la validez de esas intuiciones; porque los ejemplos que las contradicen, que acabamos de mencionar, podrán tratarse entonces como casos especiales, como casos de auto-engaño, quizá, por los que estas intuiciones pierden temporalmente su validez. Sin embargo, ninguna de estas cuestiones se va a tratar aquí. Lo que voy a hacer a continuación es criticar los argumentos del sentido común, los argumentos filosóficos y, especialmente, los argumentos científicos que se han usado como fundamento del determinismo.

Mi problema central consiste en examinar la validez de los argumentos en favor de lo que yo llamo determinismo *«científico»;* es decir, la doctrina de que la estructura del mundo es tal que *cualquier suceso puede ser racionalmente predicho, con cualquier grado de precisión que se desee, si contamos con una descripción suficientemente precisa de los sucesos pasados junto con todas las leyes de la naturaleza.*

El problema es pertinente principalmente porque los expo-

[1] Esta parte del *Post Scriptum* puede considerarse una versión perfeccionada de mi artículo «Indeterminism in Quantum Physics and in Classical Physics», *British Journal for the Philosophy of Science,* 1, núms. 2, pp. 117-133, y 3, pp. 173-195. [Sobre esta cuestión, ver también Popper, *Sobre nubes y relojes. Aproximación al problema de la racionalidad y libertad del hombre* (1966), reimpreso en *Conocimiento objetivo (Obj. Kn.,* 1972), 1974, pp. 193-235. *Ed.*]

nentes de la teoría cuántica presentan con frecuencia la situación como sigue. La física clásica, dicen, entraña lo que yo llamo determinismo «científico»; y sólo la teoría cuántica nos fuerza a abandonar la física clásica y, con ella, el determinismo «científico». Por el contrario, yo trato de mostrar que ni siquiera la validez de la física clásica nos impondría una doctrina determinista sobre el mundo.

Al criticar el determinismo me ocuparé de una teoría de las ciencias físicas y biológicas que mantuvieron los físicos, prácticamente sin excepción, hasta 1927, y Einstein, parece ser, casi hasta su muerte, en 1955 [2]. También la mantuvieron filósofos como

[2] De hecho, en 1954, Einstein parecía haber cambiado de opinión, de una manera fundamental, con respecto al determinismo. Así, se ha publicado ahora una carta de Wolfgang Pauli a Max Born, fechada el 31 de mayo de 1954, en la que Pauli habla de conversaciones mantenidas con Einstein en Princeton: «En particular, Einstein no considera el concepto de 'determinismo' tan fundamental como frecuentemente se piensa que es (como me dijo con énfasis repetidas veces...). De la misma manera, *discute* que él utilice, como crítico para la admisibilidad de una teoría, la pregunta: '¿Es rigurosamente determinista?' El punto de partida de Einstein es 'realista' más que 'determinista', lo que significa que su prejuicio filosófico es otro distinto.» Véase *The Born-Einstein Letters* (Nueva York, 1971), p. 221, o *Albert Einstein-Hedwig und Max Born: Briefwechsel, 1916-1955* (Munich, 1969), página 293.

Popper había visitado a Einstein en Princeton en 1950, y dio una conferencia en aquella ocasión (a la que asistió Einstein) con el tema de su propio ensayo sobre «El indeterminismo en la física cuántica y en la física clásica» («Indeterminism in Quantum Physics and in Classical Physics», *British Journal for the Phylosophy of Science*, 1, pp. 117-133 y 173-195) que acabaría siendo la base para este volumen del *Post Scriptum*. En su autobiografía intelectual, Popper escribe sobre las tres conversaciones con Einstein en aquella ocasión: «Nuestro tema de conversación principal fue el indeterminismo. Yo intenté persuadirle de que abandonara su determinismo, que conducía a la concepción de que el mundo era un universo cerrado parmenídeo de cuatro dimensiones, en el que el cambio era una ilusión humana, o algo muy parecido. (El corroboró que ésta había sido su concepción, y mientras discutíamos le llamé 'Parménides'.) Yo argumenté que, si los hombres, u otros organismos, podían experimentar el cambio y la sucesión genuina en el tiempo, entonces es que éste era real. No podría darse razón de él mediante una teoría de apariciones sucesivas en nuestra conciencia de láminas de tiempo que coexisten en algún sentido; porque este 'aparecer en la conciencia' tendría precisamente el mismo carácter que el de la sucesión de cambios que la teoría trata de justificar. Recurrí también a los, en cierta manera, obvios argumentos biológicos: que la evolución de la vida, y el modo en que los organismos se comportan, especialmente los animales superiores, no puede ser realmente entendida sobre la base de ninguna teoría que interprete el tiempo como si fuera algo semejante a otro espacio (anisótropo) coordinado. Después de todo, *no* experimentamos espacios coordinados. Y ello es debido a que tales espacios son simplemente no-existentes: hemos de guardarnos de hipostasiarlos, pues se trata de construcciones que son casi totalmente arbitrarias. ¿Cómo, si no, podríamos entonces experimentar el tiempo coordinado —el tiempo apropiado, efectivamente, a nuestro sistema inercial— no sólo como

Spinoza, Hobbes, Hume, Kant, Schopenhauer, J. S. Mill y (al menos hasta 1927) M. Schlick. Schlick seguía indeciso sobre la cuestión en 1930, como lo muestra la interesante cita que sigue:

> Puesto que esta suposición de que *todos* los sucesos están sometidos a las leyes universales se describe, normalmente, como el principio de la causalidad universal, puedo expresar [mi tesis] de esta manera: Toda ciencia presupone el principio de la causalidad universal...
> Toda la experiencia sustenta la creencia de que esta presuposición es satisfecha, al menos en la medida en que es necesaria en la vida práctica, en los contactos con otros hombres y con la naturaleza, y también para la suma precisión que requiere la tecnología. Pero si el principio de causalidad es válido absolutamente y en todas partes, sin la más ligera excepción —es decir, si el determinismo es correcto—, eso no lo sabemos. Lo que sí sabemos, sin embargo, es esto: que es imposible decidir sobre la cuestión del determinismo o del indeterminismo simplemente pensando y especulando, pensando el número de argumentos a favor y en contra (todos los cuales no serían más que pseudo-argumentos). Tal empresa está destinada a crear una impresión ridícula, especialmente si se piensa en el abrumador contingente

---

algo real, sino también como absoluto, es decir, como algo inalterable e independiente de cualquier cosa que podamos hacer (excepto cambiar nuestro estado de moción)?

»*La realidad del tiempo y del cambio* me parecía ser el punto esencial del realismo. (Aún lo considero así, y así ha sido considerado por algunos oponentes idealistas del realismo, tales como Schrödinger y Gödel.)

»Cuando visité a Einstein acababa de publicarse el volumen de Schilpp, *Einstein*, en *The Library of Living Philosophers*; este volumen contenía una contribución de Gödel, famosa ahora, que empleaba, contra la realidad del tiempo y del cambio, argumentos extraídos de las dos teorías de la relatividad de Einstein. Einstein se había manifestado, firmemente, a favor del realismo en ese volumen. Y disentía claramente del idealismo de Gödel: en su réplica sugería que las soluciones de Gödel de las ecuaciones cosmológicas deberían haber 'sido excluidas sobre la base de fundamentos físicos'.

»Traté entonces de presentar a Einstein-Parménides, de la manera más rigurosa posible, mi convicción de la necesidad de adoptar una actitud rotunda contra cualquier concepción idealista del tiempo. Y también traté de mostrar que, aunque la concepción idealista era compatible tanto con el determinismo como con el indeterminismo, habría que tomar una postura clara a favor de un universo 'abierto', un universo en el que el futuro no estuviera, en sentido alguno, contenido en el pasado o en el presente, aun cuando éstos imponen severas restricciones sobre aquél. Y argüí que no deberíamos permitir que la influencia de nuestras teorías nos llevase a renunciar, con demasiada facilidad, al sentido común. Era claro que Einstein no deseaba renunciar al realismo (que encontraba en el sentido común los argumentos más sólidos en su favor), aunque pienso que, como yo, estaba dispuesto a admitir que algún día podríamos vernos forzados a renunciar a él si se aducían en su contra argumentos muy poderosos (digamos, del tipo de los de Gödel).»

Véase K. R. Popper, *Búsqueda sin término* [*U. Q.*, 1976], pp. 172-174, o P. A. Schilpp, ed., *The Philosophy of Karl Popper*, 1974, vol. 1, p. 102. Véase también el apartado 26 de este mismo libro. *Ed.*

> de artificio lógico y experimental con el que la física contemporánea está cerrando cautelosamente el cerco para atacar el problema de si el principio de causalidad es válido incluso para los super-finos sucesos dentro de los átomos [3].

He citado este pasaje porque, en muchos aspectos, es representativo de la teoría que yo pretendo criticar: que el principio de causalidad es equivalente a la tesis del determinismo, y que sabemos por experiencia que es correcto, al menos para todos los fines prácticos: para «*todos* los sucesos», aunque quizá no para «'absolutamente' todos los sucesos», «en todas las partes y sin la más ligera excepción»; una duda suscitada por la teoría cuántica. También porque pretendo aceptar el desafío inherente en la sugerencia de Schlick de que el problema es discutible, pero sólo con argumentos empíricos, mientras que el pensamiento especulativo en estas cuestiones no puede usar más que «pseudo-argumentos», destinados a crear una «impresión ridícula». En realidad, yo no creo que perfeccionar un poco los argumentos empíricos exceda las posibilidades del pensamiento especulativo.

1. DETERMINISMO: RELIGIOSO, «CIENTIFICO» Y METAFISICO

La idea intuitiva de determinismo puede resumirse diciendo que el mundo es como una película de cine: la fotografía o la escena que está proyectándose es el *presente*. Las partes de la película que ya se han proyectado constituyen el *pasado*. Y las que aún no se han proyectado constituyen el *futuro*.

En la película, el futuro coexiste con el pasado; y el futuro está fijado, exactamente, en el mismo sentido que el pasado. Aunque el espectador no conozca el pasado, todo suceso futuro, sin excepción, podría en principio conocerse con certeza, exactamente

---

[3] M. Schlick, *Fragen der Ethik* (1930), p.106 (la traducción [al inglés] es mía; las cursivas son de Schlick. *Cf.* también el ensayo de Schlick en *Naturwissenschaften*, 8, 1920, esp. p. 467). Es un poco sorprendente encontrar que la filosofía determinista y, con ella, nuestro problema están siendo rechazados ahora —por cierto, por un reciente defensor de Schlick— como «un fantasma *del siglo XVIII*» (*Mind*, 1954, p. 331). El tiempo pasa, no hay duda, y no hay duda de que el tiempo, con tiempo, resolverá todos nuestros problemas. Sin embargo, es bastante raro que nosotros, los vejestorios, que recordamos los tiempos y los problemas de Einstein y de Schlick, tengamos que hacer un gran esfuerzo para recordar que esos hombres crearon sus fantasmas antes, incluso, de que Laplace produjese el suyo (al que llamó «una inteligencia») en el *Essai* de 1819.

como el pasado, puesto que existe en el mismo sentido en el que existe el pasado. De hecho, el futuro es conocido para el productor de la película, para el Creador del mundo.

La idea del determinismo tiene un origen religioso, aunque hay grandes religiones que creen en el indeterminismo: la doctrina de que al menos *algunos* acontecimientos no están fijados de antemano. (Desde San Agustín, por lo menos, la teología cristiana ha enseñado, en la mayoría de los casos, la doctrina del indeterminismo; las grandes excepciones son Lutero y Calvino.) El determinismo religioso está relacionado con las ideas de divina omnipotencia —poder total para determinar el futuro— y divina omnisciencia, que entraña que el futuro es conocido por Dios ahora, y, por tanto, cognoscible de antemano y fijado de antemano [4].

Además del determinismo religioso, existe una forma de la doctrina a la que llamaré *«científica»*.

Históricamente, se puede considerar la idea de un determinismo «científico» como el resultado de substituir la idea de Dios por la idea de naturaleza, y la idea de ley divina por la de ley natural. La naturaleza, o quizá «la ley de la naturaleza», es omnipotente y omnisciente. Todo lo fija de antemano. Al contrario que Dios, que es inescrutable, y a quien sólo puede conocerse a través de la revelación, las leyes de la naturaleza pueden ser descubiertas por la razón humana ayudada por la experiencia humana. Y si conocemos las leyes de la naturaleza podemos predecir el futuro a partir de los datos presentes por métodos puramente racionales.

Es característico de todas las formas de la doctrina determinista que *todo* suceso en el mundo esté predeterminado: si hay un solo suceso (futuro) que no esté predeterminado, hay que rechazar el determinismo, y el *indeterminismo* es verdadero. Respecto a lo que yo llamo determinismo *científico,* esto significa que, si hubiera un solo suceso futuro en el mundo que no pudiera predecirse, en principio, por medio de cálculo basado en las leyes naturales y en los datos que conciernen al estado presente o pasado del

[4] La doctrina de la omnipotencia divina trae dificultades, sin embargo. Algunas de ellas son internas o lógicas (la omnipotencia ¿incluye el poder de cambiar el pasado, o sólo implica poder sobre el futuro?). Choca también, hasta cierto punto, con la doctrina de la omnisciencia divina. (Porque si Dios lo conoce todo, entonces conoce el futuro; por tanto, está fijado de antemano y es inalterable, incluso por el propio Dios.) No discutiré aquí las dificultades éticas que entraña la doctrina de la omnipotencia divina, tales como el problema ético de si no es malo enseñar la adulación al poder.

mundo, entonces habría que rechazar el determinismo «científico».

Así, la idea fundamental del determinismo «científico» es que la estructura del mundo es tal que todo suceso futuro puede, en principio, ser calculado racionalmente de antemano sólo con que conozcamos las leyes de la naturaleza y el estado presente o pasado del mundo. Pero, si *todo* suceso ha de ser *predictible, tiene que ser predictible con cualquiera que sea el grado de precisión deseado:* porque puede alegarse que la más mínima diferencia en la medida distingue a dos sucesos *diferentes.*

Aunque la idea de determinismo «científico» parece, históricamente hablando, una especie de transferencia del determinismo religioso a términos naturalistas y racionalistas, es posible, desde luego, ver la idea de determinismo «científico» bajo un aspecto diferente. Se puede presentar, por ejemplo, como resultado de una crítica un tanto sofisticada de la concepción del mundo según el sentido común, según la cual todos los sucesos del mundo pueden dividirse en dos tipos: los sucesos predictibles, tales como el cambio de las estaciones, los movimientos diarios y anuales del sol y de las estrellas fijas, o el funcionamiento de un *reloj;* y los sucesos impredictibles, tales como los caprichos del tiempo o el comportamiento de las *nubes* [5].

Ahora bien, esta concepción, según el sentido común, de la diferencia entre *relojes* y *nubes* puede criticarse haciendo la pregunta un tanto sofisticada de si estos dos tipos de sucesos son realmente diferentes o si sólo el insatisfactorio estado de nuestros conocimientos los hace aparecer como diferentes; *si no sería el comportamiento de las nubes tan predictible como el de los relojes, si supiéramos tanto sobre nubes como sabemos sobre relojes.*

Esta pregunta, o, mejor, esta conjetura, se convirtió en convicción tan pronto como el avance del conocimiento científico hizo posible predecir los movimientos de los planetas o «vagabundos» —que fueron una vez símbolos notorios del capricho— con tanta precisión como los de las propias estrellas fijas. Fue este éxito, el éxito de las leyes de Kepler y de la dinámica de los cielos de Newton, el que llevó a la aceptación casi universal del determinismo «científico» en los tiempos modernos.

---

[5] [Véase K. R. Popper: «Sobre nubes y relojes», *op. cit.* Cuando Popper leyó esta nota y releyó esta página, se dio cuenta de que tenía olvidado que había utilizado por primera vez la imagen de las nubes y los relojes en este apartado del *Post Scriptum;* por tanto, no usó esta sección al escribir... «sobre nubes y relojes». De modo que puede haber alguna discrepancia entre las dos exposiciones. *Ed.*]

El poder de la creencia en el determinismo «científico» puede calibrarse por el hecho de que Kant, que por razones morales rechazaba el determinismo, se sintiera, no obstante, impelido a aceptarlo como un hecho innegable, establecido por la ciencia. Esto condujo a una antinomia en su sistema filosófico que nunca consiguió resolver a su entera satisfacción. Como trataré de mostrar, la mecánica de Newton no entraña, sin embargo, determinismo «científico». Si estoy en lo cierto en esto, entonces Kant estaba equivocado al creer que al aceptar la mecánica de Newton se había comprometido a aceptar el determinismo «científico», y su no resuelta antinomia, simplemente, no se plantea.

La discusión crítica del determinismo «científico» será nuestra principal tarea. Pero además del determinismo religioso, y del «científico», hay una tercera versión de la doctrina determinista, que tendrá que discutirse también, aunque sea brevemente. Puede describirse como *determinismo metafísico* [6].

La doctrina metafísica del determinismo afirma sencillamente que todos los sucesos de este mundo son fijos, o inalterables, o predeterminados. No afirma que sean conocidos por nadie; son predictibles por métodos científicos. Pero afirma que el futuro es tan inmutable como el pasado. Todos sabemos lo que quiere decir que no se puede cambiar el pasado. Es en este mismo sentido, precisamente, en el que el futuro no puede cambiarse, según el determinismo metafísico.

El determinismo metafísico, claramente, no es contrastable. Porque incluso aunque el mundo nos sorprendiera constantemente y no diera ninguna señal de predeterminación ni de regularidad siquiera, el futuro podría seguir estando predeterminado e incluso ser conocido para aquellos capaces de leer un libro del destino. El indeterminismo metafísico es igualmente incontrastable. Porque incluso aunque el mundo tuviera una apariencia totalmente regular y determinista, ello no establecería que no existiera ningún suceso no determinado de ningún tipo. Ahora bien, la falta de contrastabilidad o de contenido empírico es indicativa de debilidad lógica (no de falta de sentido, naturalmente). Una doctrina puede ser demasiado débil lógicamente como para ser contrastada. Y,

[6] [* (Añadido en 1981). Evité, o traté de evitar, el término «ontología» en este libro, igual que en los otros; especialmente por el alboroto que organizan algunos filósofos sobre «ontología». Quizá hubiera sido preferible explicar este término, y luego usarlo, a evitarlo. Sea lo que fuere, las cuestiones de terminología no son nunca importantes.]

por la misma razón, puede ser implicada por otra doctrina. Así, el determinismo metafísico, por su debilidad, es implicado tanto por el determinismo religioso como por el «científico», y puede decirse de él que contiene sólo lo que es común a las diversas teorías deterministas. Es irrefutable sólo a causa de su debilidad. Pero esto no significa que los argumentos en favor o en contra de él sean imposibles. Los argumentos más fuertes a su favor son los que sostienen el determinismo «científico». Si fallan, queda poco que pueda sostener el determinismo metafísico. Por esta razón voy a ocuparme de ellos en primer lugar. Sólo hacia el final de mi exposición (en la sección 26) propondré algunos argumentos más directos en contra del determinismo metafísico.

En las secciones siguientes se examinarán, primero, los argumentos del sentido común y los argumentos filosóficos en favor del determinismo «científico»; sólo después me ocuparé de los argumentos basados en la física clásica.

2. PREGUNTAS DE-POR-QUE. CAUSALIDAD Y DETERMINISMO «CIENTIFICO»

En la sección anterior, he sugerido una distinción del sentido común o *prima facie* entre sucesos predictibles e impredictibles, o entre *relojes* y *nubes,* y he sugerido también que puede considerarse que el determinismo «científico» surge de una crítica sofisticada de esta distinción. Otra forma en la que el sentido común aborda la idea del determinismo es por medio de la idea popular de causalidad. Uno de los argumentos más sencillos y plausibles en favor del determinismo es éste: *siempre* podemos preguntar, de *todo* suceso, por qué ocurrió, y de toda pregunta tal de-por-qué, siempre podemos obtener, en principio, una respuesta que nos ilumine. Así, todo suceso es «causado»; y esto parece significar que debe estar determinado, de antemano, por los sucesos que constituyen su causa.

Podríamos admitir la verdad de esas consideraciones sin admitir, sin embargo, que nos conducen a la idea del determinismo «científico». De hecho, tiene un interés considerable determinar con precisión el lugar donde terminan los argumentos de sentido común y el lugar donde empieza la sofisticación.

Consideremos, en primer lugar, algunas preguntas de-por-qué típicas y algunas de las respuestas típicas que se les dan y que son perfectamente satisfactorias en el plano del sentido común.

«¿Por qué almacenan miel las abejas?» Respuesta: «Porque la necesitan para comer durante el invierno.» (Esto difícilmente puede considerarse un enunciado teórico.) «¿Por qué hay eclipse de Luna hoy?» Respuesta: «Porque hoy la Tierra está entre el Sol y la Luna, de forma que su sombra cae sobre la Luna.» (Este no es todavía un enunciado teórico a partir del cual podamos predecir un eclipse.) «¿Por qué murió?» Respuesta: «Porque cuando asistió a un funeral, la semana pasada, estuvo bajo la lluvia cerca de media hora. Así que cogió un catarro que se convirtió en neumonía, y, después de todo, tenía setenta y tres años.» (Muchos han sobrevivido la experiencia incluso a los setenta y cuatro años.)

Podemos suponer que todas esas respuestas son perfectamente aceptables, y que dan precisamente el tipo de información que se pide. Incluso los niños, que repiten sus preguntas de-por-qué insaciablemente, no piden generalmente que se les dé una explicación «mejor», en el sentido de una que les permita *predecir* sucesos del tipo en cuestión. Normalmente, lo que quieren que se les dé es un enunciado de otras causas más, causas que preceden en el tiempo a las que ya se han dado o que las motivarían más aún: los niños no repiten sus preguntas porque quieran una serie más completa de condiciones, sino porque quieren que se les dé una «cadena causal» de sucesos consecutivos.

El hecho de que siempre podamos hacer preguntas de-por-qué, y de que siempre podamos obtener respuestas pertinentes, no tiene, pues, en sí mismo, nada que ver con el determinismo, sea «científico» o de otro tipo.

Pero ahora podemos avanzar un paso más y exigir que las respuestas a nuestras preguntas de-por-qué —es decir, las explicaciones ofrecidas— consistan efectivamente en condiciones iniciales (causas) a partir de las cuales los hechos a explicar puedan ser deducidos lógicamente, si se conocen las leyes universales pertinentes [7]. Este, ciertamente, es un paso que se aleja de la teoría popular de la causalidad, y que se acerca a una teoría más sofisticada. Aceptemos esta exigencia, y supongamos, además, una «ley de la causalidad universal» con el fin de que *todo* suceso pueda, en principio, explicarse causalmente en el sentido de nuestra exigen-

[7] *Cf.* Sección 12 de *La lógica de la investigación científica [L. Sc. D.]*

cia; es decir, suponemos que siempre hay «causas» (condiciones iniciales) y leyes universales que nos permiten deducir el «suceso» en cuestión. Esta es una suposición fuerte. Pero todavía no llega a determinismo «científico», y esto por varias razones.

Primero, la idea del sentido común de un «suceso» (que ha de explicarse causalmente) es principalmente cualitativa. Así, la exigencia del determinismo «científico» de que debemos ser capaces de predecir el acontecimiento *con cualquiera que sea el grado de precisión deseado* va, ciertamente, más allá de la idea del sentido común de la causalidad universal. Puede ser de sentido común pedir una explicación causal de la fiebre de Juan; pero queda fuera del sentido común pedir una explicación de por qué su temperatura está entre 39,1° y 39,2° en vez de entre 39° y 40°, o pedir una predicción con el correspondiente grado de precisión.

Segundo, la idea del sentido común de causa es igualmente cualitativa. La comprensión de que las causas —es decir, las condiciones iniciales— no se nos dan nunca con precisión absoluta; de que, por tanto, tenemos que conformarnos con condiciones iniciales que son hasta cierto punto imprecisas, y de que este hecho plantea sus propios problemas peculiares: todo esto va más allá de la idea del sentido común o intuitiva de causalidad.

Tercero, está el siguiente problema que surge de los dos anteriores: el determinismo «científico» exige la capacidad de predecir todo suceso con cualquiera que sea el grado deseado de precisión, siempre que se nos den unas condiciones iniciales *suficientemente* precisas. Pero ¿qué significa aquí «suficientemente»? Obviamente, tendremos que explicar «suficientemente» de tal manera que nos privemos a nosotros mismos del derecho a argüir —cada vez que fallemos en nuestras predicciones— que no se nos dieron unas condiciones iniciales suficientemente precisas.

En otras palabras, *nuestra teoría tendrá que dar razón* de la imprecisión de la predicción. Dado el grado de precisión que exigimos a la predicción, tendrá que permitirnos calcular el grado de precisión de las condiciones iniciales que sería suficiente para darnos una predicción con el grado exigido de precisión. Llamo a esta exigencia «el principio de poder dar razón». Esta exigencia tendrá que incorporarse a la definición de determinismo «científico».

Puesto que los tres puntos mencionados aquí —y especialmente el tercero— van claramente más allá de la idea popular, o del sentido común, de causalidad, es concebible que la idea popular,

intuitiva, de causalidad —incluso en la forma establecida del principio de causación universal— sea válida, hasta donde alcanza, mientras que, al mismo tiempo, la doctrina del determinismo «científico» no es válida.

Así, tenemos que precavernos contra el error que cometen tantos filósofos, que creen que es válido argumentar en favor del determinismo señalando que cada suceso tiene una causa.

3. EL PRINCIPIO DE PODER DAR RAZON

En principio, el resultado de un cálculo no será más preciso que el menos preciso de sus datos, y, en consecuencia, una predicción no será, en principio, más precisa que ninguna de las condiciones iniciales dadas sobre las que se basa [8]. Si, por tanto, exigimos que sea siempre posible que nuestras predicciones sean tan precisas como deseamos, no será posible, en principio, satisfacer esta exigencia, a menos que podamos incrementar la precisión de las condiciones iniciales pertinentes tanto como deseemos: las condiciones iniciales tendrán que ser *suficientemente precisas* para obtener la solución del problema planteado por la tarea de predicción.

Para el objetivo de definir el determinismo «científico», la exigencia de ser capaces de obtener predicciones con un grado estipulado de precisión, siempre que se nos den condiciones iniciales *«suficientemente precisas»*, sería, claramente, demasiado vaga. Una exigencia tal haría trivial nuestra definición. Siempre podríamos mantener que está satisfecha, incluso aunque siempre fallásemos en las predicciones; siempre podríamos disculpar nuestros fallos alegando que las condiciones iniciales no eran «suficientemente precisas». Para remediar esta situación, tenemos que exigir que sea posible descubrir, *antes de* contrastar el resultado de nuestras predicciones, si las condiciones iniciales son suficientemente precisas o no; dicho de otro modo, tenemos que poder determinar de antemano, a partir de la tarea de predicción (que debe enunciar,

[8] Esta regla es aplicable a la mayoría de los casos, pero no es universalmente válida: no es necesariamente aplicable, por ejemplo, a mecanismos capaces sólo de estados discretos, tales como un armonio o una máquina de escribir: nuestra información sobre la posición *precisa* del dedo del organista en una tecla determinada, por ejemplo, puede ser bastante irrelevante para el efecto predicho. El lenguaje verbal, especialmente si está escrito o impreso, tiene esencialmente este carácter, y lo mismo ocurre con las computadoras digitales.

entre otras cosas, el grado de precisión que se exige de la predicción) junto con la teoría, cómo tienen que ser de precisas las condiciones iniciales o «datos» para que nos permitan realizar esa tarea concreta de predicción. Para expresarlo más completamente, debemos *poder dar razón de antemano* de cualquier fallo en predecir un suceso con el grado deseado de precisión, señalando que nuestras condiciones iniciales no son bastante precisas y estableciendo lo precisas que tendrían que ser para esa tarea de predicción. Así, cualquier definición satisfactoria de determinismo «científico» tendrá que estar basada en el principio (principio de poder dar razón) de que *podemos calcular a partir de nuestra tarea de predicción (junto con nuestras teorías, naturalmente) el grado requerido de precisión de las condiciones iniciales* [9].

Algunas tareas de predicción son «capaces de dar razón»; es decir, satisfacen el principio de poder dar razón mientras que otros pueden no ser capaces de ello. También podemos decir de una teoría que es «capaz de dar razón» cuando las tareas de predicción dentro de ella son, por regla general, capaces de dar razón.

Para ciertos propósitos puede ser útil operar con un principio de poder dar razón algo más fuerte, y a él se llega haciendo referencia a la *precisión de los resultados de las mediciones posibles a partir de los cuales pueden calcularse las condiciones iniciales,* en lugar de hacer referencia a la *precisión de las condiciones iniciales.* Así pues, en este sentido más fuerte, una tarea de predicción puede no ser capaz de dar razón porque no podemos determinar a partir de ella (y de la teoría) el grado requerido de precisión de las *mediciones* posibles sobre las que vayamos a basar nuestras predicciones. Sin embargo, es concebible que la misma tarea de predicción sea capaz de dar razón en el sentido más débil de permitirnos calcular el grado de precisión con el cual tendrían que ser dadas las *condiciones iniciales* para resolverla.

Esta idea más fuerte de la capacidad de dar razón es claramente la más «realista» de las dos: una teoría que fuera capaz de dar razón en el sentido más débil, pero no en el más fuerte, sería una

---

[9] Un principio similar fue formulado por Pierre Duhem, pero en un contexto diferente (no en una discusión sobre determinismo) y con un propósito diferente. El lo expresa como sigue: «Es necesario que seamos capaces de determinar... el error que puede permitirseles a los datos si queremos obtener un resultado dentro de un grado definido de aproximación.» *Cf. The Aim and Structure of Physical Theory* (1954), p. 143, el tercer párrafo a partir del final del capítulo III.

teoría cuyo carácter determinista no podríamos contrastar, en principio: no podría usarse en apoyo del determinismo «científico». Dicho de otro modo, el determinismo «científico» requiere la capacidad de dar razón en el sentido más fuerte. No obstante, de ahora en adelante estaré pensando siempre en la capacidad de dar razón en el sentido más débil, a menos que aluda especialmente a la diferencia entre los dos sentidos. La razón es que si una teoría es *incapaz* de dar razón en el sentido más débil (de «capaz de dar razón»), entonces, claramente *no* es capaz de dar razón tampoco en el más fuerte; en otras palabras, *incapacidad* de dar razón, en el sentido más débil de la frase «capaz de dar razón», entraña (o es, lógicamente hablando, más fuerte que) *incapacidad* de dar razón en el sentido más fuerte de la frase «capaz de dar razón».

Puesto que el determinismo «científico» entraña el principio de poder dar razón, cualquier ejemplo de un problema de predicción decididamente incapaz de dar razón aplicable a nuestro propio mundo destruiría al instante la doctrina del determinismo «científico». Pero, incluso aunque no podamos presentar un ejemplo indudable de este tipo, debe quedar claro que *no tenemos ninguna razón para creer en el determinismo «científico» si no tenemos ninguna razón para creer que el principio de poder dar razón resulta universalmente satisfactorio.*

En las secciones siguientes trataré de mostrar que, al menos en dos campos, algunos argumentos convincentes del sentido común, al igual que algunos famosos argumentos filosóficos en favor del determinismo, se desbaratan precisamente porque no tenemos razón alguna para creer que el principio de poder dar razón está satisfecho en esos campos.

Nuestra exposición establecerá lo siguiente: es posible que continuemos aprendiendo cada vez más sobre un campo determinado, que podamos predecir cada vez más sucesos y que podamos continuar incrementando la precisión de nuestras predicciones. Y, sin embargo, este crecimiento continuo de nuestra capacidad de predicción puede no constituir una razón válida para la creencia de que el determinismo «científico» es válido dentro de ese campo. Nuestras predicciones pueden mejorar constantemente y, sin embargo, al mismo tiempo, pueden ser obtenidas por métodos que ni siquiera sugieran que el principio de poder dar razón está satisfecho.

Más adelante (en la sección 17) se planteará el problema de la capacidad de dar razón de la física clásica, y se mostrará que ape-

nas hay razones para creer que es capaz de dar razón en el sentido más débil, y que hay muy buenas razones para creer que no es capaz de dar razón en el sentido más fuerte.

4. EL ESTUDIO DE LA CONDUCTA Y EL PRINCIPIO DE PODER DAR RAZON

Examinemos ahora un argumento muy sencillo, y a la vez muy fuerte, en favor del determinismo.

Los indeterministas han afirmado, a veces, que los hombres, y quizá en menor grado los animales superiores, se comportan de forma muy distinta a como se comportan los sistemas planetarios o los relojes mecánicos, y que el determinismo (incluso si es válido en el campo de la mecánica) puede no ser válido en el campo de la biología. En contra de esa opinión, algunos deterministas han argumentado como sigue:

Es innegable que podemos predecir a menudo la conducta de los animales, y también de las personas, con mucho éxito. Además, esas predicciones tienden a ser mejores según vamos aprendiendo más sobre el hombre y sobre el animal; y pueden perfeccionarse aún más mediante el estudio sistemático de su conducta. No hay razón por la que este proceso de aprender cada vez más sobre la conducta tenga que llegar a su fin. Así, el resultado de nuestro estudio de los organismos puede llegar a ser el mismo que el de los sistemas planetarios. Podemos expresar esto diciendo que los organismos superiores pertenecen a la categoría de los relojes. (No importa si descubrimos que son como relojes mecánicos o, más bien, como un mecanismo electrónico autorregulado.) Podemos dejar abierta, de momento, la cuestión de si las *nubes* pertenecen también a esta categoría.

Este argumento —*«el argumento desde el estudio de la conducta»*, como puede llamarse— es bueno como sentido común y, en mi opinión, muy revelador. Pero no logra sus objetivos; ni siquiera si concedemos, como yo estoy dispuesto a hacer, que no hay límite al posible perfeccionamiento, por medio de un estudio cada vez más profundo, de nuestras predicciones sobre la conducta animal y humana. Como tentativa de apoyo del determinismo, el argumento desde el estudio de la conducta, sencillamente, no es válido.

Para mostrarlo, sólo necesitamos hacer referencia al principio

de poder dar razón. El determinismo «científico» no afirma simplemente que podamos mejorar nuestras predicciones mediante el aumento de nuestro conocimiento, sino que exige *que seamos capaces de calcular a partir de nuestra tarea de predicción específica el grado de precisión de la información inicial que nos es necesaria para realizar la tarea de predicción.*

Ahora bien, el argumento desde el estudio del comportamiento no contiene nada que sugiera que el perfeccionamiento de nuestro conocimiento nos ayudará a satisfacer este principio. Resulta que yo tengo bastante éxito cuando predigo lo que va a hacer mi gato a continuación: si va a saltar sobre mi mesa de despacho y tumbarse en mi cuaderno, o si va a saltar al alféizar de la ventana y de allí al jardín: y continuamente aprendo más cosas sobre su conducta. Pero la conducta sobre la que aprendo consiste, principalmente, en a) acciones significativas (u «orientadas hacia un fin»), o b) hábitos, o formas de hacer las cosas. Estudiar los últimos puede ayudar a completar algunos detalles de los esquemas más amplios de las primeras. No obstante, cuando predigo que va a sentarse en mi cuaderno, hay muchos detalles que no puedo predecir. Por ejemplo, *puedo fácilmente equivocarme en varias pulgadas.*

No hay nada en la situación de mi aprendizaje que me sugiera cómo podría añadir esos detalles al cuadro. Naturalmente, siempre podemos decir que un mejor conocimiento de las condiciones iniciales pertinentes reduciría las pulgadas tanto como quisiéramos. *Pero simplemente no sabemos qué clase de condiciones iniciales pueden ser pertinentes a la tarea* de predicción de esas pulgadas. No es sólo que no tengamos una teoría de la conducta que satisfaga el principio de poder dar razón: hasta ahora no tenemos ni idea de dónde buscar tal teoría.

Puede objetarse que un estudio preciso del sistema nervioso, especialmente del cerebro, haría mucho por terminar con las lagunas en nuestras predicciones. Esto puede ser cierto, por lo que yo sé, y lo aceptaré aquí por cuestiones de argumentación. Pero significa *renunciar al argumento desde el estudio de la conducta.* Sustituye el argumento del sentido común de que podemos aprender cada vez más del estudio de la conducta de los animales por otro argumento totalmente diferente: el argumento de que la fisiología y la física son sistemas deterministas.

5. TEMPERATURAS CRITICAS Y EL PRINCIPIO DE TODO-O-NADA

El hecho es que sabemos tan poco de esas cuestiones que no tenemos ni la más ligera idea de cómo aplicar nuestro muy considerable conocimiento de fisiología del cerebro a una tarea de predicción como la que concierne a la posición *precisa* de mi gato.

Pero supongamos que sabemos cómo aplicar nuestro conocimiento de fisiología del cerebro. Supongamos más especialmente que lo que necesitamos son condiciones iniciales que nos permitan predecir la contracción de un músculo determinado; y, en último caso, condiciones iniciales que nos permitan predecir si un ganglio (o grupo de ganglios) efectuará una descarga.

Ahora bien, el proceso de una descarga nerviosa es análogo en muchos aspectos a una explosión: el nervio se descarga repentinamente cuando cierto potencial eléctrico (el potencial de la placa terminal) se ha elevado a una cierta altura crítica. Si no se alcanza esa altura, el nervio no se descarga [10]. (Esto se conoce como el principio de todo-o-nada de la transmisión nerviosa.) De manera similar, una explosión química sólo tendrá lugar si se alcanza una cierta temperatura crítica, y puede ocurrir que por debajo de ella no ocurra nada [11].

Hay, sin embargo, graves dudas sobre si el principio de poder dar razón es aplicable a la temperatura crítica de una explosión, y precisamente, por razones análogas, sobre si es aplicable a la transmisión nerviosa. Admitamos que, si la temperatura está por debajo de la crítica y se va elevando lenta pero constantemente, podremos predecir el momento de la explosión. Sin embargo, la capacidad de dar razón entraña la tesis de que seamos capaces de predecir el momento de la explosión con tanta exactitud como queramos. Y esta tesis, a su vez, supone otra: que seamos capaces de tomar la medida de la temperatura y de su tasa de incremento con tanta exactitud como queramos. Pero la temperatura es un efecto-masa; es una cantidad molar o macroscópica. Es esencialmente una media; y, en principio, las cantidades como ésa no son susceptibles de medirse con una precisión estipulada.

---

[10] Véase, por ejemplo, J. C. Eccles, *The Neurophysiological Basis of Mind*, 1953. [Véase también K. R. Popper y J. C. Eccles: *El yo y su cerebro (S. & I. B., 1977) Ed.*]

[11] Véase, por ejemplo, K. F. Bonhoeffer, «Modèles Physico-Chemiques de l'Excitation Nerveuse», *Journal de Chemie Physique*, 51, pp. 521-9.

Hay muchas razones para creer que el valor preciso del potencial que requiere un nervio dado para efectuar una descarga depende también de otros efectos-masa. Por ejemplo, depende de un efecto-fatiga (que, muy probablemente, depende a su vez de la presencia o ausencia de la suficiente *concentración*, es decir, de *un número suficientemente grande* de moléculas de cierto tipo). Incluso aunque supongamos la verdad del determinismo metafísico, no hay nada en nuestras teorías actuales que indique que vayamos a ser capaces alguna vez de calcular, en general, la precisión necesaria de las condiciones iniciales a partir de las especificaciones de la tarea de predicción.

Para resumir estas consideraciones, podemos decir que el cerebro (tanto si amplifica procesos cuánticos elementales, como si no) es, con toda probabilidad, altamente sensible a efectos-masa, tales como temperatura o concentración de ciertos productos químicos. (Esto está lejos de ser sorprendente, si consideramos que nuestros movimientos musculares son efectos-masa y que dependen, en parte, de otros efectos-masa tales como «descargas» de los impulsos de los nervios.) Pero no tenemos clave, ni indicación de ninguna clase, que nos permita decir que el principio de poder dar razón puede aplicarse a esos efectos-masa, o, en caso afirmativo, de qué manera puede aplicarse.

Vemos lo lejos que estamos ya del «argumento desde la conducta» del sentido común, y que, en efecto, nada en nuestra experiencia normal de aprender cada vez más sobre la conducta sugiere que el principio de poder dar razón vaya a ser satisfecho.

Podemos decir, en general, que aunque nuestro conocimiento, y con él nuestra capacidad de predecir, puede aumentar constantemente en un cierto campo, este hecho no puede usarse nunca por sí mismo como argumento en favor de la opinión de que algo parecido al determinismo «científico» es sostenible en ese campo. Porque nuestro conocimiento puede aumentar constantemente sin acercarse a esa clase especial de conocimiento que satisface el principio de poder dar razón.

6. RELOJES Y NUBES

Puesto que he tenido que referirme a los efectos-masa, éste es quizá un buen lugar para criticar el argumento determinista de los relojes y las nubes, mencionado más arriba en la sección 1. Estoy

refiriéndome al argumento de que la distinción del sentido común entre sucesos predictibles (el movimiento de los planetas o de los relojes) y los sucesos impredictibles (el tiempo o el movimiento de las nubes) no es válido y de que esta distinción desaparecerá cuando obtengamos tanto *conocimiento* sobre las nubes —sobre las leyes que las rigen, tanto como sobre sus condiciones iniciales— como tenemos sobre los relojes.

Como demuestra el ejemplo de los planetas, hay algo en favor de este argumento. Un conocimiento más perfecto y más detallado sobre las nubes puede, en efecto, hacer mucho por asemejarlas a los sucesos de la categoría de los relojes. Pero esta semejanza no puede tener un éxito completo, porque en los sucesos relacionados con las nubes se incluyen efectos-masa.

Además, podemos ver la cuestión también desde el extremo opuesto: si deseamos predecir el funcionamiento de los relojes cada vez con mayor detalle, tendremos que investigar, por ejemplo, el flujo del calor en el reloj (para descubrir, por ejemplo, cómo afecta a la longitud del péndulo). Pero este estudio más detallado llevaría claramente a asemejar el reloj a una nube de moléculas, sujetas a movimientos de una clase que presenta los mismos problemas de predicción que los de una nube. Así, el convincente argumento de que con el aumento de nuestro saber la categoría de las nubes se aproximará cada vez más a la de los relojes puede refutarse señalando que también es argüible lo contrario.

Se dice con frecuencia que dos relojes que sean exactos el uno al otro darán la misma hora y continuarán haciéndolo. Puede que sea verdad, pero no tiene gran interés, porque nunca tenemos dos relojes que sean exactos entre sí. Es más, dos relojes producidos en serie en una fábrica y que parezcan perfectamente exactos el uno al otro no continuarán, en general, dando la misma hora. Esa es la razón por la cual están hechos de manera que puedan ajustarse con ayuda de un mecanismo regulador. Después de ajustarlos, puede que sigan dando la misma hora bastante bien; pero en general ya no serán tan iguales: una parte importante mecánicamente —el mecanismo regulador— puede que exhiba ahora una diferencia claramente observable. Esta diferencia hubo de introducirse con el fin de que los dos relojes se parecieran más en otro aspecto: en el de dar la hora. Esto muestra que la semejanza superficial puede ser muy engañosa.

Si un reloj se retrasa, un buen relojero puede encontrar la cau-

sa: quizá una mota de polvo en el mecanismo. El ejemplo es interesante, porque, aunque está de acuerdo con la ley de la causalidad universal, es claro que no satisface el principio de poder dar razón. Ningún relojero podría predecir, al ver la mota de polvo, que iba a causar un retraso de tres minutos diarios en lugar de cinco. Tampoco podría predecir que, una vez removida esa mota de polvo concreta, el reloj iba a continuar funcionando sin necesidad de mayores ajustes.

7. ARGUMENTOS DESDE LA PSICOLOGIA EN FAVOR DEL DETERMINISMO

Los argumentos que vienen de la conducta de la psicología no son válidos, como hemos visto. La razón no es tanto que haya dificultades para medir, por ejemplo, las condiciones iniciales de la conducta; es, más bien, que la precisión de nuestras predicciones de la conducta no pueden perfeccionarse indefinidamente por medio del perfeccionamiento de la precisión de las mediciones de las condiciones iniciales de la conducta.

Pero el argumento de la conducta no es más que una nueva versión de un argumento más antiguo, *el argumento desde la psicología.*

Algunos grandes filósofos han usado este argumento para refutar la idea del libre albedrío, y con ello, indirectamente, apoyar la idea del determinismo. Este argumento es ya muy antiguo. Desde el renacimiento, ha sido quizá formulado con la mayor claridad por Hobbes, Spinoza, Hume y Priestley [12]. Como dice Hobbes, «la voluntad es también necesariamente *causada* por otras cosas sobre las que no decide», y, por tanto, «se desprende que las acciones voluntarias tienen todas ellas causas necesarias».

Hume introdujo la idea de *«inferencia de motivos* a acciones voluntarias», como lo expresa; «Parece —escribe— que la conexión entre motivos y acciones voluntarias es tan regular y uniforme como cualquier parte de la naturaleza». Además, introdujo la idea de referencia «del carácter a la conducta».

---

[12] Hobbes, *Quaestiones de libertate et necessitate contra Doctorem Bramhallum* (1656); Spinoza, *Ethics*, parte I, prop. 32, corolario 2; parte II, prop. 49, escolio (véase, en especial, la respuesta a la cuarta objeción); Hume, *Treatise*, 1739, libro II, secciones I a III, «On Liberty and Necessity» *(Essays Moral, Political, and Literary*, 1742); Priestley, *The Doctrine of Philosophical Necessity* (1777).

Kant aceptó ambas ideas, y mantuvo que una información psicológica completa nos permitiría «calcular de antemano, y con certeza, de la misma forma que lo hacemos con los eclipses solares o lunares, la conducta futura de cualquier hombre» [13]. Esta afirmación muestra lo firmemente que creía en el determinismo «científico». La cuestión es si está respaldado por las observaciones psicológicas, tal como pensaba Hume, sobre la conjunción regular entre motivos y acciones.

Ahora bien, yo no quiero negar la existencia de causas psicológicas —de deseos, esperanzas, motivos e intenciones—, o de la sensación de que una situación exige una acción determinada. Por el contrario, creo que hay muchos casos importantes en los que las explicaciones causales fisiológicas y físicas están destinadas al fracaso [14], mientras que, al menos, algo parecido a una explicación psicológica satisfactoria puede darse (o quizá una desde el punto de vista de «la lógica de la situación» [15]).

Hay, por ejemplo, buenas razones para creer que ciertos procesos de aprendizaje —el aprender a andar, a hablar, a escribir, a esquiar, a tocar el piano— comprenden cambios fisiológicos tales como, quizá, la creación de circuitos nerviosos característicos; también, al practicar lo que hemos aprendido —al andar o al hablar— tienen lugar reacciones fisiológicas características. Pero no hay nada que respalde la opinión de que si un esquiador *inventa* una *nueva* secuencia de movimientos, o un escritor una *nueva* secuencia de palabras, o un compositor una *nueva* secuencia de acordes, los procesos fisiológicos vayan a ser diferentes de aquellos en los que las mismas secuencias de palabras o de acordes fueron tomadas (inconscientemente) de otros. Desde luego, la invención de algo nuevo va a veces acompañada de una sensación de excitación que puede tener alguna correlación fisiológica; pero, en otros casos, el innovador puede no ser consciente de su innovación e incapaz de distinguir entre lo que ha aprendido de otros y lo que ha creado él mismo.

---

[13] *Critique of Practical Reason*, ediciones 4.ª a 6.ª, p. 172; o WW, ed. Cassirer, vol. 5, p. 108. El pasaje está citado de forma más completa en la sección 16, más abajo.

[14] *Cf.* el último párrafo de mi artículo «El lenguaje y el problema del cuerpo y la mente», *Proceedings of the XIth International Congress of Philosophy*, vol. VII, páginas 101 y ss. [Reimpreso en *El desarrollo del conocimiento científico: Conjeturas y refutaciones* (*C. & R.*, 1963), capítulo 12.]

[15] Véase mi *Miseria del historicismo* [*Poverty*, 1957], sección 31, y *La Sociedad abierta y sus enemigos* [*O. S.*, 1945], capítulo 14.

Es importante darse cuenta, en relación con esto, de que el hecho de que cierto logro tenga el carácter de novedad es una cuestión de interpretación y de evaluación. Tomemos un ejemplo extremo: alguien puede aplicar una cita muy gastada a una situación que hace su uso no sólo inesperado, sino también creativo e ingenioso.

Oportunidad, ironía y originalidad son, pues, ejemplos de cualidades que pueden caracterizar la expresión de una persona y que no se puede esperar que se expliquen por medio de la *fisiología.* Y, sin embargo, puede haber explicaciones *psicológicas* de la más diversa índole: la influencia de la atmósfera familiar; ejemplos y contactos que pueden haber llevado a establecer normas inconscientes; el hecho de leer y escribir y otras influencias de la educación pueden usarse para explicar tales cosas. La invención de una nueva demostración matemática encuentra aquí su explicación, igual que la invención de nuevos argumentos, y quizá incluso el intento de continuar un argumento.

Así pues, hay pocas razones para creer, y muchas para no creer, que un fisiólogo pueda predecir, partiendo del estudio del cerebro de un matemático, los pasos de una nueva demostración que el matemático está a punto de inventar. Pero, una vez que la demostración está ante nuestros ojos, podemos encontrar alguna explicación *psicológica* para algunos de sus pasos: uno puede parecerse a una demostración previa del mismo matemático; otro puede ser una aplicación ingeniosa de un método inventado por su maestro. Un análisis más penetrante puede incluso revelar algunas de las causas que le hicieron pensar, en esta situación determinada, en su antiguo profesor y en el tipo de consejos que había recibido de él.

En todos estos casos, nuestra explicación causal puede utilizar, en parte o totalmente, hipótesis psicológicas que pueden ser contrastadas por medio de predicciones que han tenido éxito. Así, podríamos haber predicho con éxito que Euler, enfrentado a un problema de dificultad moderada, sería capaz de resolverlo; o que Mozart compondría una misa o una ópera, si hubiera aceptado el encargo de hacerlo, y que, considerando tanto su habilidad como su conciencia, compondría una gran obra y no simplemente algo hecho de cualquier manera para salir del paso. No hay razón alguna para no ir más lejos y formular hipótesis psicológicas contrastables y ya contrastadas sobre las cuales se puedan basar predicciones, incluso predicciones relativas a nuevas invenciones.

Estas consideraciones sobre las «causas» psicológicas exponen —lo admito— una dificultad fundamental en la formulación intuitiva de la doctrina del «libre albedrío». Difícilmente encontraremos una decisión, una acción o una nueva invención que sea «incausada» en el sentido de que no podamos «explicarla causalmente» de una forma satisfactoria, siempre que sepamos mucho, no sólo del ambiente, los deseos, los temores y las actitudes de la persona en cuestión, sino también de la manera en que sus decisiones pueden resultar influenciables por argumentos o por sus gustos musicales o literarios. Así, si consideramos que la doctrina del «libre albedrío» trata de afirmar las «decisiones sin causa» en este sentido, las críticas de Hobbes, Spinoza, Hume y sus sucesores tienen que admitirse como correctas [16].

Ninguna de esas consideraciones puede usarse, sin embargo, como argumento en favor de nada parecido al determinismo «científico». Porque el determinismo «científico» afirma mucho más que la existencia de causas. Afirma (como indica mi cita de Kant) que esas causas nos permiten predecir un suceso con el grado de precisión que se desee. Por tanto, implica el principio de poder dar razón; es decir, la posibilidad de calcular, partiendo de la tarea de predicción, el grado de precisión con el cual las «causas» —es decir, las condiciones iniciales— deben ser conocidas para poder resolver el problema de predicción.

Pero no hay razón alguna para creer que este principio puede satisfacerse en el campo de la psicología del aprendizaje, o de la motivación, mejor de lo que se satisface en el campo de la conducta o en el de la fisiología. Por el contrario, parece que hay mucha razón para conjeturar que no será satisfecho nunca en el campo de la psicología.

La sugerencia de Kant, por ejemplo, está claramente basada en un malentendido. Está basada en la creencia errónea de que, si podemos perfeccionar indefinidamente nuestro conocimiento de las «causas» en un campo determinado, podremos hacer nuestras predicciones en ese campo tan precisas como queramos. Pero es claro que esto no es cierto en general, ni siquiera plausible.

---

[16] Estoy suponiendo aquí que consideraciones lógicas, tales como, por ejemplo, nuestra percepción de que dos teorías son interdeducibles, o quizá contradictorias, pueden considerarse como una «causa»; ver mi «El lenguaje y el problema del cuerpo y la mente», *op. cit.*, esp. secciones 6.1 a 6.4. [Véase también la exposición de Popper sobre la interacción del «Mundo 3» con los «Mundos 1 y 2» en *El yo y su cerebro (S. & I. B.), Ed.*]

La idea de predecir, por métodos psicológicos, la acción de un hombre con el grado deseado de precisión es, desde luego, tan completamente ajena a todo el pensamiento psicológico que es difícil comprender lo que entrañaría. Entrañaría, por ejemplo, predecir, con cualquiera que fuese el grado deseado de precisión, a qué velocidad subiría un hombre las escaleras si esperase encontrar una carta anunciándole un ascenso en el trabajo, o que le habían despedido. Esto comprendería la combinación de condiciones iniciales físicas de todo tipo (la altura de las escaleras, el rozamiento de los zapatos y las escaleras) y de condiciones iniciales fisiológicas (el estado general de salud del hombre, de su corazón, de sus pulmones, etc.) con, por ejemplo, condiciones iniciales económicas (los ahorros de ese hombre, sus oportunidades de encontrar otro empleo, el número de personas dependientes de él) y condiciones iniciales psicológicas (su confianza en sí mismo o su ansiedad, etc.). Nadie sabe cómo podrían estimarse todas esas cosas, incluso aunque pudieran conocerse; cómo tendrían que evaluarse, y, especialmente, cómo tendrían que usarse las condiciones psicológicas, de tal forma que pudieran tratarse de la misma manera que las fuerzas físicas con las que habrían de ser comparadas y combinadas.

Un psicoanalista puede, en muchos años de estudio (no pocos análisis duran más de una década), desenterrar todo tipo de «causas» —motivos y qué sé yo— enterradas en el inconsciente de su paciente. Supongamos que el análisis es capaz, en gran número de casos, de predecir la conducta de su paciente con éxito. Aun así, pocos creerán, a pesar de todo su conocimiento sobre los motivos de su paciente, que el analista es capaz de predecir el tiempo exacto que le tomará a su paciente, bajo diferentes condiciones psicológicas, subir las escaleras. El analista quizá pueda decir que podría realizar la predicción si contase con los datos suficientes. *Pero no será capaz de establecer y dar razón de qué datos serían suficientes para su propósito.*

El conocimiento psicológico de un hombre (o de un gato) puede permitirnos predecir que no cometerá un asesinato o un robo (o que el gato no morderá o arañará). Pero para establecer el determinismo «científico» haría falta mucho más.

Una vez que comprendamos lo que el determinismo «científico» implica —y especialmente el principio de poder dar razón— vemos que el conocimiento psicológico tendría que ser complementado también por conocimientos fisiológicos, exactamente de

la misma manera que el conocimiento de la conducta tiene que ser complementado con conocimientos fisiológicos (como vimos cuando hablamos del argumento desde la conducta). Esto significa, naturalmente, que el argumento desde la psicología se derrumba.

Ni que decir tiene que el argumento desde la psicología era, desde el principio, más vulnerable que el de la conducta. No tanto, creo yo, porque no podamos medir la intensidad de los motivos, ya que las mediciones del conductista no le sirven, como hemos visto, como porque el uso de un concepto tal como «motivo» o «carácter» es, como nos muestra un porqué de reflexión, poco más que un torpe intento de encontrar conexiones que parezcan leyes, o incluso de inventarlas cuando no las encontramos. Yo no niego que una pregunta como «¿Cuál fue el motivo de su acción?» o una pregunta de-por-qué tal como «¿Por qué lo hizo?» sean perfectamente razonables; y también es razonable una respuesta como: «Lo hizo por celos (o por ambición, o por venganza).» Pero todas las respuestas de este tipo, incluso cuando son muy sofisticadas, no son más que rudos intentos de clasificar; o, en el mejor de los casos, de construir un esquema situacional hipotético [17] que haga la acción racionalmente inteligible. Son, en toda la extensión, tentativas para entender *post hoc;* esto es así incluso en las raras ocasiones en que operamos con un esquema que podría ser contrastado confrontándolo con predicciones.

8. LA VISION DETERMINISTA DEL MUNDO

Ni los argumentos desde la conducta ni los argumentos desde la psicología se basan en la experiencia: pocas personas afirmarían haber hecho predicciones precisas en esos campso. A mí me parecen, más bien, brotar de la convicción previa de que *el mundo físico* es determinista. En un mudo físico determinista está claro que no hay espacio para un comportamiento indeterminista; porque todo comportamiento consiste en sucesos dentro del mundo físico. Por otro lado, podría parecer que hay espacios para estados de conciencia que no están determinados. Pero el supuesto de que existen tales estados de conciencia sería muy insatisfactorio y, en realidad, gratuito. No podrían tener ningún tipo de cone-

[17] Ver las referencias de las notas 14 y 15 más arriba.

xión causal con la conducta. No podríamos saber de ellos, o, en todo caso, seríamos incapaces de hablar de ellos. Porque si lo hiciéramos, sucesos indeterminados tendrían cierta influencia causal sobre el mundo físico de los sonidos; un supuesto que contradiría la doctrina de que el mundo físico es determinista.

De este modo, una concepción determinista del mundo físico nos impone una concepción determinista del mundo de la conducta y del mundo de la psicología; y, en efecto, la concepción del mundo físico de Hobes y sus sucesores era determinista. Esto no prueba, naturalmente, que el argumento desde la psicología que ellos expusieron (o el argumento desde la conducta que expusieron Hume y Schlick) fuese consecuencia de esta convicción; pero el argumento mismo nos sugiere que es consecuencia de ella.

Es interesante que la creencia de Hobbles de que el mundo físico era determinista precedió a la teoría de Newton. El magnífico éxito de Newton pudo entonces ser fácilmente interpretado como una impresionante corroboración de la doctrina determinista. Parecía que Newton hubiera convertido el antiguo programa determinista en una realidad.

Esto explica la absoluta confianza en la verdad del determinismo que encontramos, por ejemplo, en Kant.

La sugerencia de que los argumentos desde la conducta y desde la psicología son consecuencia de una concepción determinista del mundo físico explicaría también por qué ninguno de los que propusieron esos argumentos se paró a considerar el problema de la capacidad de dar razón. Porque en una visión del mundo físico hobbesiana, según la cual el mundo funciona como un reloj, la capacidad de dar razón parece intuitivamente obvia [18]; y lo mismo ocurre con la visión newtoniana (al menos si no profundizamos demasiado en los detalles del problema de-muchos-cuerpos). Si el mundo físico fuera determinista y si el principio de poder dar razón fuera satisfecho en el campo de la física, entonces no habría necesidad de preocuparse por él en los campos de la conducta y de la psicología.

Creo que nuestra exposición muestra que los argumentos populares o del sentido común en favor del determinismo, igual que los argumentos filosóficos tradicionales, no son válidos. Pero también sugiere que podemos esperar que los argumentos más fuertes

[18] El intento de Descartes de hacer sitio, en su concepción del mundo como un mecanismo de relojería, para mentes indeterminadas que interfieren con la materia, no es convincente.

provengan de la física clásica. Antes de volverme a la física, sin embargo, voy a explicar de forma general algunas razones por las que el indeterminismo debería considerarse aceptable *prima facie* y por qué la carga de la prueba debería descansar sobre el determinista.

9. LA CARGA DE LA PRUEBA

Una razón importante para aceptar el indeterminismo, al menos a título de ensayo, es que la carga de la prueba recae sobre los hombros del determinista. Los únicos argumentos razonables fuertes que yo conozco en favor del determinismo son aquellos que sustentan el determinismo «científico»; y, en vista del hecho de que los argumentos del sentido común en favor del determinismo científico parecen desbaratarse cuando se los enfrenta al problema de la capacidad de dar razón, da la impresión de que la balanza del sentido común se inclina a favor del indeterminismo.

Hay varias razones por las que me parece que la carga de la prueba recae sobre el determinista. Mencionaré sólo cuatro.

Primero, el nada sofisticado sentido común apoya la opinión de que hay relojes *y* nubes, es decir, sucesos que son más predictibles y sucesos que lo son menos; que predeterminación y predictibilidad son cuestiones de grado.

Segundo, hay un argumento *prima facie* en favor de la opinión de que los organismos están menos predeterminados y son menos predictibles que, al menos, algunos de los sistemas más simples, y que los organismos superiores están menos predeterminados y son menos predictibles que los inferiores.

Los castores (y los hombres) producen cambios característicos e inconfundibles en su entorno físico. Sin duda, el entorno físico puede corresponder produciendo cambios característicos e inconfundibles en los castores (o en los hombres). Pero para establecer la verdad del determinismo habría que demostrar muchas más cosas. Suponiendo que los castores no han existido siempre, el determinista tendría que demostrar que las condiciones físicas (distintas de la presencia de los castores) pueden producir, de una forma predictible, castores. Pero, aunque sabemos mucho sobre las condiciones físicas producidas por los castores, no sabemos nada, ciertamente, sobre las condiciones físicas que pueden producir castores. Aquí hay una asimetría en nuestro saber, y la carga

de probar que la laguna en nuestro conocimiento puede rellenarse recae sobre el determinista. Hasta ahora no tiene más que un programa, en el mejor de los casos.

La tercera razón, y la que yo considero más importante, está estrechamente relacionada con el problema de la «libre voluntad». Si el determinismo es verdad, debería ser posible, en principio, que el físico o el fisiólogo que no sabe nada de música predijera, al estudiar el cerebro de Mozart, los lugares del papel en los que Mozart iba a poner su pluma. Aún más, el físico y el fisiólogo deberían ser capaces de anticiparse a la acción de Mozart y escribir su sinfonía antes, incluso, de que Mozart la concibiera conscientemente. Lo mismo debería ocurrir con los descubrimientos matemáticos, y todas las demás contribuciones a nuestro saber. A pesar del hecho de que un hombre como Kant estuvo implícitamente comprometido a sostenerlas, estas conclusiones me parecen, intuitivamente, absurdas. Absurdas o no, van mucho más allá que nada de lo que nosotros sabemos; así, una vez más, la carga de la prueba recae sobre el determinista.

En cuarto lugar, el indeterminismo, que afirma que existe, *al menos,* un suceso que no está predeterminado y no es predictible, es claramente una aserción más débil que el determinismo «científico», que afirma que todos los sucesos son, en principio, predictibles. Aunque yo prefiero, en la ciencia, las teorías fuertes a las débiles, lo hago porque son más discutibles, es decir, criticables. En todo caso, aquel que propone la teoría más fuerte acepta la carga de la prueba: tiene que dar argumentos en favor de su teoría, principalmente exhibiendo la capacidad explicatoria de dicha teoría. Pero el determinismo, incluso la versión que yo llamo «científica», no pertenece a la ciencia, ni tiene capacidad explicatoria alguna.

CAPITULO II

DETERMINISMO «CIENTIFICO»

10. EL DETERMINISMO *PRIMA FACIE* DE LA FISICA CLASICA. EL DEMONIO DE LAPLACE

Los físicos cuánticos dicen a menudo que lo que ellos llaman «física clásica» (que comprende las teorías de Newton, de Maxwell e incluso las de Einstein) implica determinismo, mientras que la física cuántica implica indeterminismo. Sin admitir que esa afirmación sea cierta, estoy dispuesto, desde luego, a admitir que hay una diferencia entre la física clásica y la física cuántica. La teoría cuántica es una teoría de probabilidades, mientras que la física clásica tiene un carácter diferente [1]. Me propongo descubrir la física clásica como «determinista *prima facie*», indicando con esta denominación que no quiero prejuzgar la cuestión de si entraña o no algún tipo de determinismo.

Lo que yo llamo «el carácter determinista *prima facie*» de la física clásica puede describirse mejor con la ayuda del llamado «demonio de Laplace».

Laplace creía que el mundo consistía en corpúsculos que actuaban unos sobre otros según la dinámica de Newton, y que un conocimiento completo y preciso del estado inicial del sistema del mundo en un instante del tiempo sería suficiente para deducir su estado en cualquier otro instante. (El «estado» de un sistema newtoniano está dado cuando están dadas las condiciones iniciales completas; es decir, las posiciones, masas, velocidades y direcciones del movimiento de todas sus partículas [2].) Un conocimiento de este tipo es claramente sobrehumano. Por eso Laplace introdujo la ficción de un demonio: una inteligencia sobrehumana, capaz de averiguar el conjunto completo de las condiciones iniciales

---

[1] Hay teorías que no encajan en ninguna de esas dos categorías, especialmente las teorías cualitativas y clasificatorias.

[2] Respecto a las teorías de campo, debería entenderse que las «condiciones iniciales» del sistema (en el sentido que yo empleo) incluyen las condiciones de los límites.

del sistema del mundo en cualquier instante del tiempo. Con la ayuda de esas condiciones iniciales y con las leyes de la naturaleza, es decir, las ecuaciones de la mecánica, el demonio sería capaz, según Laplace, de deducir todos los estados futuros del sistema del mundo; esto mostraría que, siempre que se conociesen las leyes de la naturaleza, el futuro del mundo estaría implícito en cualquier instante de su pasado; y así quedaría establecida la verdad del determinismo [3].

El punto crucial del argumento de Laplace es éste: *Hace de la doctrina del determinismo una verdad de la ciencia y no de la religión.* El demonio de Laplace no es un Dios omnisciente, sino simplemente un supercientífico. No se le supone capaz de nada que no puedan hacer los científicos humanos, o al menos que no puedan hacer aproximadamente: simplemente se le supone capaz de realizar sus tareas con perfección sobrehumana.

Así, Laplace hubiera estado dispuesto a admitir que los científicos humanos no pueden averiguar las condiciones iniciales de todos los cuerpos físicos del universo; pero hubiera señalado que pueden medir todas las condiciones iniciales de un sistema solar, si el número de planetas es reducido. También hubiera estado dispuesto a admitir que los científicos no pueden obtener condiciones iniciales, absolutamente exactas; pero hubiera señalado que pueden perfeccionar el grado de precisión al medirlas y que no hay límite absoluto a ese perfeccionamiento. Una vez más, Laplace hubiera admitido que, si el sistema contiene más de dos cuerpos, la teoría de Newton nos permite calcular el estado futuro del sistema sólo por un método de aproximación, de acuerdo con el estado actual de nuestros conocimientos matemáticos; pero Laplace hubiera señalado que, aunque no hemos resuelto el problema general de —muchos— cuerpos —es decir, el problema de calcular la interacción gravitatoria newtoniana de más de dos cuerpos—, podríamos encontrar la solución algún día, lo que ciertamente legitimaría el haber invertido al demonio con este conocimiento; y podría haber añadido que, aunque el problema general pudiera ser estrictamente insoluble, podemos, en cada caso *particular* (siempre que no sea demasiado complejo), substituir la solución exacta por una aproximación que tenga el grado de precisión que decidamos estipular.

[3] [Véase P. S. Laplace: *Essai philosophique sur les probabilités,* 1819, Introducción. Véase también Popper: *El yo y su cerebro —S. & I. B.—. Ed.*]

Es en este sentido en el que el demonio de Laplace no es más que un científico humano idealizado. En realidad es un Laplace idealizado. Laplace creía que había resuelto el gran problema de la estabilidad de nuestro sistema solar. Creía que había demostrado —bajo el supuesto de que el sistema era cerrado, es decir, que ningún cuerpo extraño podía entrar o interferir en él desde el exterior— que los planetas continuarían, *en todos los períodos futuros de tiempo,* manteniendo sus distancias medias del Sol. (Veremos en las secciones 12 y 14 que estaba equivocado.)

Se supone que el demonio de Laplace opera, como un científico humano, con condiciones iniciales y con *teorías,* es decir, sistemas de leyes naturales. Las teorías que, para sistemas físicos apropiados, respondan plenamente al propósito de Laplace pueden ser denominadas «deterministas *prima facie*».

Esta denominación se introduce aquí para caracterizar ciertos rasgos de las teorías de Newton, Maxwell o Einstein en contraposición con otras teorías, como ocurre con la termodinámica, o la mecánica estadística, o la teoría cuántica y quizá también la teoría de los genes. Yo sugiero la siguiente definición:

Una teoría física es determinista *prima facie* si, y sólo si, nos permite deducir, a partir de una descripción *matemáticamente exacta* del estado inicial de un sistema físico cerrado que se describe en términos de la teoría, la descripción, *con cualquiera que sea el grado finito de precisión estipulado,* del estado del sistema en cualquier instante dado del futuro.

Esta definición no exige predicciones matemáticas exactas, aunque supone que las condiciones iniciales son, matemáticamente, absolutamente exactas. No podemos exigir más, si queremos estar seguros de que la mecánica de Newton no queda excluida por la definición, puesto que sólo se conocen métodos aproximados para la solución del problema de más de dos cuerpos.

Puede argüirse, incluso, que, por razones similares, deberíamos debilitar nuestra definición añadiendo las palabras «siempre que el sistema en cuestión no sea demasiado complejo». Porque no se sabe si existen métodos satisfactorios para resolver por aproximación problemas de —muchos— cuerpos, cuando el sistema en cuestión contiene muchísimos cuerpos, y especialmente si las masas y las distancias de los cuerpos son todas del mismo orden de magnitud. La cuestión es, naturalmente, que, incluso con condiciones iniciales matemáticamente exactas, los métodos de compu-

tación numérica introducen sus propias imprecisiones que, en el caso de algunos sistemas complejos, podemos ser incapaces de reducir, por medio de pasos sucesivos de aproximación, por debajo de cierto nivel [4]. Así, puede resultar imposible obtener una predicción con el grado de precisión deseado.

Aunque este punto es importante, no continuaré con él aquí; por el contrario, *supondré* —como concesión a mis oponentes deterministas— que las teorías de Newton y Maxwell son deterministas *prima facie* en el sentido de mi definición original.

Una vez adoptada esta definición, la cuestión que tenemos delante es la siguiente: suponiendo que sea *verdadera* una teoría física determinista *prima facie*, ¿estamos autorizados a inferir de dicha suposición la verdad del determinismo «científico»? o, dicho de otro modo, ¿estamos autorizados a inferir el carácter determinista del mundo del carácter determinista *prima facie* de una teoría?

Más adelante (en la sección 13) daré varias razones por las que creo que esa inferencia no sería válida. Pero mi próxima tarea es explicar con mayor claridad la idea de determinismo «científico».

11. LA IDEA DE DETERMINISMO «CIENTIFICO»: PREDICTIBILIDAD DESDE DENTRO

La idea general de determinismo puede explicarse, como hemos visto, con la ayuda de la metáfora de una película que muestra los estados sucesivos del mundo. En esta película, lo que se mostrará en el futuro está tan fijado, o determinado, como lo que se ha mostrado en el pasado, y, puesto que el futuro está fijado, podría, en principio, ser conocido de antemano; no simplemente imaginado, sino conocido con certidumbre.

Teniendo en cuenta esta metáfora, podríamos decir que el determinismo «científico» es consecuencia del intento de sustituir la vaga idea de conocimiento anticipado del futuro por la idea más precisa de *predictibilidad de acuerdo con los procedimientos científicos racionales de predicción.* Es decir, el determinismo afirma que el futuro puede *deducirse racionalmente* a partir de las condiciones iniciales pasadas o presentes en unión de teorías universales verdaderas.

[4] Véase también J. von Neumann y H. H. Goldstine: «Numerical Inverting of Matrices of High Order», *Bull. Am. Math. Soc.*, 53, 1947, pp. 1022-99.

Es importante percatarse de que el determinismo «científico» afirma más que la mera posibilidad, o incluso la existencia, de conocimiento anticipado del futuro, y que, por tanto, se presta más fácilmente a las críticas. Sería clara y lógicamente posible que todo suceso fuera conocido de antemano por alguien (que lo soñase, por ejemplo), incluso en un mundo en el que muchos sucesos ocurrieran de una manera fortuita, y no estuvieran sujetos a nada parecido a leyes universales. En un mundo así, el determinismo «científico» sería falso, puesto que no habría teorías verdaderas suficientemente fuertes como para usarlas como base de una predicción científica racional. Podemos expresar esto diciendo que el determinismo «científico» afirma más que el carácter de película del mundo. Afirma que los sucesos de la película no son nunca fortuitos, sino que siempre están sujetos a normas, de forma que cada fotograma o escena de la película nos permite *calcular por métodos racionales* cualquiera de los fotogramas o escenas que siguen, con ayuda de las normas o leyes que conectan los sucesivos fotogramas. A menos que implique todo esto, una doctrina determinista no será de tipo «científico».

Lo dicho puede, quizá, clarificar la relación entre la idea general de determinismo y esa versión de él que yo he denominado «científica». El punto crucial es que el último apela al *éxito de la ciencia humana,* tal como la teoría newtoniana: el determinismo «científico» debe aparecer como resultado del éxito de la ciencia empírica o, al menos, apoyado por ella. Parece estar basado en *la experiencia humana.*

Por esto fue, sin duda, por lo que Laplace apeló no a un Dios omnisciente, sino simplemente a un demonio investido con poderes que, pretendía Laplace, no sobrepasaran *en principio* a los del científico humano. Laplace no suponía que su demonio conociese ningún estado futuro del mundo por intuición; eso no idealizaría los poderes de la razón humana, sino que, en principio, iría más allá de ellos. Sin embargo, Laplace suponía que su demonio conocía las condiciones iniciales con precisión, sin duda porque pensaba que no había límites al posible perfeccionamiento del saber humano en esta dirección. De la misma manera, suponía que su demonio conocía la condición de *todos* los corpúsculos en un sistema, cualquiera que fuese su grado de complejidad; de nuevo, sin duda, porque pensaba que no había límite posible al conocimiento humano en este aspecto, aunque se daba cuenta de que ningún ser humano (ni ninguna computadora de ningún tamaño

finito) podría nunca, en la práctica, hallar las coordenadas de todos los corpúsculos en un sistema como un gas, que contiene un número muy grande de ellos. Podemos, quizá, expresar mejor la intención de Laplace diciendo que los poderes del demonio excederían a los del científico humano sólo en cantidad; serían ilimitados sólo en los campos en los que no existiera *ningún límite definido* a los poderes del científico humano.

Esta idea puede precisarse más con la ayuda de los requisitos importantes con los que Laplace hubiera estado de acuerdo, sin duda. Esos dos requisitos, formulados aquí haciendo referencia al demonio, serán más adelante, en una forma ligeramente más abstracta, incorporados a nuestra definición final de determinismo «científico».

El primer requisito es el que sigue:

1) Al demonio, igual que a un científico humano, *no* debe suponérsele capaz de *averiguar las condiciones iniciales con precisión matemática absoluta;* como un científico humano, tendrá que conformarse con un grado finito de precisión. Pero puede suponerse que el demonio puede hacer que el margen de imprecisión de sus mediciones sea tan pequeño como quiera; es decir, más pequeño que cualquier margen finito que cualquiera pueda especificar.

A primera vista, este requisito no parece más que un mínimo ajuste de nuestra definición de una teoría determinista *prima facie;* porque hemos denominado teoría determinista *prima facie* a aquella que nos permite calcular cualquier predicción con el grado de precisión deseado o especificado, si se nos dan unas condiciones iniciales *matemáticamente exactas.* La doctrina del determinismo «científico» requiere un poco más: que nosotros —o el demonio— seamos capaces de calcular cualquier predicción con el grado de precisión deseado o especificado si se nos dan unas condiciones iniciales de *un grado finito de imprecisión* (contando siempre con que la imprecisión no exceda un grado que podemos estipular de antemano sobre la base de la tarea de predicción, de acuerdo con el principio de poder dar razón).

El segundo requisito es el que sigue:

2) Debe suponerse que el demonio, como un científico humano, forma parte del mundo físico, cuyo futuro va a predecir; al menos debe suponerse que existen procesos *físicos* en el mundo

que pueden interpretarse: *a)* como procesos por los cuales el demonio puede obtener información, *b)* como procesos de cálculos de la predicción, y *c)* como procesos de formulación de la predicción. Dicho de otro modo, debemos imaginarnos al demonio no como un espíritu incorpóreo, fuera del sistema físico que va a predecir, sino más bien como la encarnación física de un espíritu; es decir, debe haber, de algún modo, una interacción entre sus actividades esenciales y el sistema. Podemos resumir este requisito diciendo que debe *predecir el sistema desde dentro de él* y no desde fuera.

Esta segunda exigencia es también derivable del requisito de que el demonio no debe estar en posesión de poderes que, *en principio*, sobrepasen todos los poderes humanos. Esta no es una exigencia mía *ad hoc*. Durante al menos treinta años, los físicos han supuesto tácitamente que esta exigencia era parte de la doctrina del determinismo. Esto resulta claro cuando se recuerdan ciertos argumentos de Heisenberg, argumentos a los que puedo referirme aquí sin aceptarlos por ello [5]. Me refiero al argumento de que el determinismo no sirve en vista del hecho de que, debido a la interferencia del proceso de medir con el estado del sistema medido, hay límites definitivos a la precisión posible de nuestro conocimiento de las condiciones iniciales y, por tanto, a las predicciones que pueden calcularse a partir de ellas. Este argumento equivale a rechazar la idea de que el demonio puede ser un espíritu incorpóreo: supone que no debe ser dotado de capacidad ilimitada allí donde, en principio, cualquier ser humano y cualquier instrumento físico encontrarían limitaciones. Dicho de otro modo, el argumento de Heisenberg contra el determinismo se basa en el supuesto implícito de que el determinismo entraña *predictibilidad desde dentro*, con el grado de precisión que se desee.

---

[5] No estoy de acuerdo con estas ideas de Heisenberg. Creo que debo decirlo aquí, con bastante dureza, puesto que en mi artículo «Indeterminism in Quantum Physics and in Classical Physics», donde formulé por primera vez lo que se presenta aquí como el requisito 2), hablé a veces como si aceptase la idea de que el indeterminismo cuántico es debido a la interferencia del proceso de medición con los objetos medidos. Esta forma de hablar se debía, supongo, al hecho de que cuando estaba escribiendo ese artículo mi objetivo no era criticar la física cuántica, sino, más bien, mostrar que ciertas características que se creen peculiares de ella son también aplicables a la física clásica. [Ver Popper: «La mecánica cuántica sin 'El Observador'», reimpreso como la introducción a *Teoría cuántica y el cisma en física*, vol. III del *Post Scriptum*. *Ed.*]

12. DOS DEFINICIONES DE DETERMINISMO «CIENTIFICO»

Podemos definir ahora el determinismo «científico» como sigue:

La doctrina del determinismo «científico» *es la doctrina que dice que el estado de cualquier sistema físico cerrado en cualquier instante futuro dado puede ser predicho, incluso desde dentro del sistema, con cualquiera que sea el grado estipulado de precisión, mediante la deducción de la predicción a partir de teorías, en conjunción con concisiones iniciales cuyo grado de precisión requerido puede calcularse siempre (de acuerdo con el principio de poder dar razón) si la tarea de predicción es dada.*

Esta es la *definición más débil* y, sin embargo, es suficientemente fuerte como para formular las ideas implicadas en un determinismo «científico».

Lo que nuestra definición exige es la predictibilidad de *cualquier* suceso mediante la exigencia (como se ha indicado en secciones anteriores) de la predictibilidad del estado de *cualquier* sistema físico, en cualquier momento futuro, con *cualquier* grado de precisión que se especifique. Además, incorpora los requisitos 1) y 2) que se explicaron en la sección anterior, y también el principio de poder dar razón. Todos ellos son ingredientes indispensables de la idea de determinismo «científico».

Por otro lado, podrían formularse definiciones *más fuertes:* hay posibles ingredientes de la idea de determinismo «científico» que algunos pueden considerar, intuitivamente, partes esenciales de él, y que son omitidos en nuestra definición. Estoy pensando, especialmente, en la idea de que podemos predecir de cualquier sistema *si un suceso de un tipo dado ocurrirá en él o no.* En otras palabras, podríamos añadir a la definición que acabamos de dar el requisito de que pueda predecirse, de cualquier estado dado, *si el sistema en cuestión estará alguna vez en ese estado o no.*

Si añadimos este requisito a nuestra definición, obtendremos lo que podemos llamar *la versión más fuerte del determinismo «científico».*

La cuestión de si ocurrirá alguna vez un eclipse de Sol (o, digamos, un eclipse de Sol seguido a los quince días por uno de Luna) es un ejemplo que pueden tener en mente aquellos que consideran *la versión más fuerte como esencial para el determinismo «científico».* Otro ejemplo, más importante para nuestra

exposición, es la cuestión estudiada por Laplace de si nuestro sistema solar es estable, o, usando una formulación más concreta, si la distancia media entre el Sol y cualquier planeta será alguna vez, por ejemplo, el doble de la distancia media presente, o su mitad.

Fue, en parte, la creencia de que había resuelto este problema lo que sugirió a Laplace la idea del demonio. Así, puede decirse que la versión más fuerte está muy cerca de la que Laplace tenía en mente.

13. ¿SE DESPRENDE EL DETERMINISMO «CIENTIFICO» DE UNA TEORIA DETERMINISTA *PRIMA FACIE*?

A primera vista puede parecer que la definición de determinismo «científico», incluso en su versión más fuerte, es tan estrechamente similar a la de una teoría determinista *prima facie* que la verdad del determinismo «científico» se desprendería inmediatamente de la verdad de cualquier teoría determinista *prima facie*, tal como la mecánica de Newton. Esta impresión explica, sin duda, por qué no sólo Kant y Laplace, sino también tantos otros grandes pensadores que creían firmemente en la verdad de la mecánica de Newton, pensaron que estaban obligados a aceptar una doctrina como el determinismo «científico». También Einstein creyó en la validez de la inferencia [6], y lo mismo hicieron sus oponentes, los defensores de la interpretación oficial de la teoría cuántica («la interpretación de Copenhague»). Y, sin embargo, la inferencia no es válida.

Primero, hay que percatarse de que hay una diferencia considerable entre lo que yo he llamado el carácter determinista *prima facie* de una teoría y el determinismo «científico». Al afirmar el primero, siempre afirmamos *de una teoría* que tiene una determinada propiedad. Al afirmar el segundo, afirmamos *del mundo* que tiene una determinada propiedad. Es cosa admitida que, si una teoría es verdadera, entonces describe ciertas propiedades del mundo. Pero esto no significa que para cada propiedad de una teoría verdadera haya la correspondiente propiedad del mundo.

Como primera indicación de que puede ser peligroso confiar

[6] [Pero véase capítulo I, nota 2, que sugiere que Einstein abandonó esta idea. *Ed.*]

en la impresión de que el determinismo «científico» se desprende del carácter determinista *prima facie* de una teoría, debería recordarse lo que sigue. Incluso si fuéramos a suponer que la mecánica de Newton es verdadera, estaría claro que él no habría obtenido una teoría que entrañase determinismo «científico», si no por otra razón, por la de que él no mostró que todos los sucesos físicos fueran mecánicos; sólo tras una deducción de éxito a partir de la mecánica de Newton, de una teoría satisfactoria de la electricidad, del magnetismo, y de la óptica, podría plantearse la cuestión de si la verdad de la mecánica de Newton puede usarse como argumento en favor del determinismo «científico» o no. O, dicho de otro modo, el determinismo «científico» podría derivarse, en caso de que pudiera hacerse, sólo de un sistema de física que fuera *completo* o *comprensivo,* en el sentido de que permitiera la predicción de todo tipo de sucesos físicos [7].

Una segunda advertencia en contra de que se confíe en la impresión de que el determinismo «científico» se desprende de la verdad de una teoría determinista *prima facie* puede derivarse del hecho de que *la versión más fuerte,* en todo caso, del determinismo «científico» es falsa, incluso si suponemos que el mundo es un sistema puramente mecánico (sin electricidad, etc.) y que la mecánica de Newton es verdadera. Esto se mostrará en la sección siguiente, con ayuda de uno de los resultados de Hadamard. Más adelante, trataré de mostrar más, no sólo que incluso la versión más débil del determinismo «científico» es incompatible con ciertas teorías —por ejemplo, con la de Einstein—, sino también que debe rechazarse por razones lógicas.

14. UN RESULTADO DE HADAMARD

En un artículo muy interesante [8], publicado en 1898, Hadamard exponía un sencillo problema mecánico: el movimiento, con velocidad constante, de un punto-masa a lo largo de las geodésicas —es decir, las líneas más rectas— de una superficie curva

[7] Debe recordarse que Kant y Laplace florecieron precisamente en el período de la historia de la ciencia (1785-1818) en que había más razones que nunca para creer que el programa de reducir toda la física a la mecánica de Newton tenía probabilidades de éxito.

[8] J. Hadamard: «Les surfaces à courbures opposées», etc. *Journal des Mathématiques pures et appliquées,* 5.ª serie, vol. 4 (1898), pp. 27-73; véase también Pierre Duhem: *The Aim and Structure of Physical Theory,* pp. 139 y ss.

infinita (de un tipo especial, esto es, con una curvatura negativa variable; se supone que no hay discontinuidades). Hadamard supone que nos dan la posición inicial (el punto de partida del movimiento) con absoluta precisión, y permite que la dirección inicial del movimiento varíe dentro de un ángulo α. Hadamard demuestra que entonces habrá varios *tipos de cursos*, especialmente (i) *órbitas* o *cursos cerrados*, incluidas curvas que estén cerradas sólo asintóticamente, de tal forma que un punto que se mueva sobre ellas estará siempre a una distancia finita del punto de partida, y (ii) *trayectorias hacia el infinito*, tales que, después de un período de tiempo suficientemente largo, un punto que se mueva sobre ellas excederá cualquier distancia finita dada del punto de partida [9]. Consideramos que dos órbitas (cursos cerrados) diferentes salen de nuestro punto de partida con dos direcciones iniciales diferentes que encierran un ángulo α pequeño. Hadamard muestra que, aunque hagamos α tan pequeño como queramos, habrá, no obstante, trayectorias *yendo al infinito* que salgan del punto de partida *dentro* del ángulo α; es decir, entre dos *órbitas* cualesquiera que escojamos.

Pero esto significa que ninguna medida de la *dirección* inicial de la trayectoria, por precisa que sea (excepto una precisión matemática absoluta), puede determinar si el punto-masa se está moviendo en una órbita o si lo hace en una trayectoria que se dirige al infinito; ni siquiera bajo el supuesto, nada realista, de que la *posición* inicial sea dada con absoluta precisión. Significa, dicho de otro modo, que no podemos determinar si el punto-masa se moverá de tal forma que su distancia del punto de partida no exceda nunca un valor finito, o si a la larga empezará a incrementar constantemente su distancia y se irá al infinito.

Así, la versión más fuerte del determinismo «científico» que se discutía en esta sección es refutada por los resultados de Hadamard. Porque, como señala Hadamard [10], ningún grado finito de

---

[9] Hadamard distingue una tercera categoría de geodésicas —los casos límites— que no nos concierne aquí.

[10] Hadamard, *op. cit.*, p. 71 (sección 59). Hadamard dice allí que sus resultados sugieren que el problema de la estabilidad del sistema solar «cesaría de tener sentido». No estoy de acuerdo. Demostrar que un problema de predicción no puede calcularse sobre la base de ninguna teoría ni de ninguna medición de las condiciones iniciales, por muy precisas que sean, no hace que el problema pierda sentido, sino que establece que es insoluble. Un problema insoluble no es un problema sin sentido y el descubrimiento de su insolubilidad puede dilucidarlo tanto como el descubrimiento de una solución.

precisión de las condiciones iniciales nos permitirá predecir si un sistema planetario (de muchos cuerpos) será estable en el sentido de Laplace, o no. Esto es debido al hecho de que los estados iniciales matemáticamente exactos que determinan las órbitas y otros que determinan las geodésicas que van al infinito no pueden ser desenredados, como hemos visto, por ninguna medición física. Por esto, Hadamard refuta el resultado de Laplace (mencionado más arriba); un resultado que muy bien puede haber sido una de las mayores inspiraciones de la idea de Laplace de determinismo «científico».

Sin embargo, Hadamard no refuta, en mi opinión, la doctrina más débil del determinismo «científico» tal como la he definido más arriba. Todavía podemos obtener predicciones *para cualquier instante dado* de tiempo del estado del punto-masa, siempre que midamos su dirección inicial con un grado de precisión que dependerá *a)* del lapso del tiempo entre la partida y el instante mencionado en la tarea de predicción, y *b)* del grado de precisión especificado en la predicción. Lo que *no podemos* predecir es el comportamiento del sistema *para todos* los instantes del tiempo [11].

[11] En esta formulación, la diferencia entre «para todo» y «para cualquier... dado» muestra una cierta analogía con el hecho (descubierto por Gödel) de que, aunque podamos construir para cualquier enunciado aritmético dado una teoría formalizada en la cual el enunciado sea resoluble, no podemos construir una teoría formalizada en la que todos los enunciados aritméticos sean resolubles. Pensando en esta analogía, podemos decir que una cuestión hadamardiana (que pregunta si un sistema de-muchos-cuerpos, cuyas condiciones iniciales sean dadas con un grado finito de precisión, estará alguna vez en un estado especificado) es una *cuestión físicamente irresoluble.*

CAPITULO III

EL ARGUMENTO EN FAVOR DEL INDETERMINISMO

15. POR QUE SOY INDETERMINISTA: LAS TEORIAS COMO REDES

Yo, personalmente, creo que la doctrina del indeterminismo es verdadera y que el determinismo carece por completo de fundamento.

Entre las razones de mi convicción destaca el argumento intuitivo (mencionado en la sección 7, más arriba) de que la creación de una obra nueva, tal como la *Sinfonía en sol menor* de Mozart, no puede predecirse en todos sus detalles por un físico o un fisiólogo que estudie con detalle el cuerpo de Mozart —especialmente su cerebro— y su entorno físico. La opinión contraria parece, intuitivamente, absurda; en cualquier caso, parece obvio que sería dificilísimo dar argumentos razonables en su favor y que no hay nada, en el momento actual, excepto un prejuicio cuasi-religioso para apoyar esa opinión o el prejuicio de que la omnisciencia de la ciencia se aproxima de alguna manera, aunque sólo sea en principio, a la omnisciencia divina.

Admito francamente que este argumento se relaciona estrechamente con el problema tradicional de la «libre voluntad» que, sin embargo, no voy a tratar. El problema que me interesa ahora es, más bien, el que plantea nuestro ejemplo sobre Mozart: si el mundo es de tal manera que podríamos, en principio, si quisiéramos lo suficiente, predecir incluso sucesos tan únicos como la creación de una nueva sinfonía en todos sus detalles, por los métodos racionales de la ciencia. Este es el único problema que me interesa en este terreno. Tengo que dejar esto bien claro, porque estoy sencillamente aburrido del análisis del significado de las palabras «libre» y «voluntad» y de la cuestión de si Mozart u otro cualquiera podría haber actuado de manera diferente de como lo hizo. Me interesa el mundo de los hechos; y, aunque desde Schlick (que introdujo el análisis de significados en este campo, bajo la in-

fluencia de Wittgenstein) se acepta ampliamente que incluso Hume se preocupó del análisis del significado de las palabras, esto es, a mi modo de ver, un error [1]. No dudo que también Hume estuviera interesado en la estructura del mundo y que aclarase malentendidos verbales sólo donde pensaba que eran obstáculos al entendimiento del mundo.

Es, pues, la supuesta predictibilidad científica de obras únicas lo que me interesa y lo que encuentro absolutamente increíble. Esta es una de esas cuestiones, algunas de las cuales se enumeraron en la sección 9, en las que la carga de la prueba recae sobre el determinista.

Pero no es suficiente cargar al determinista con la responsabilidad de proporcionar argumentos para afirmaciones frecuentemente repetidas que me parecen a mí alocadas, y para las cuales nunca se han ofrecido buenos argumentos. Hay también argumentos filosóficos contundentes, en parte lógicos y en parte metafísicos, *contra* el determinismo; argumentos que, hace muchos años, me convencieron de la debilidad del determinismo «científico» [2].

Yo considero nuestras teorías científicas como invenciones humanas, redes creadas por nosotros para atrapar el mundo. Por supuesto, éstas difieren de las invenciones de los poetas e incluso de las de los técnicos. Las teorías no son *sólo* instrumentos. A lo que aspiramos es a la verdad: contrastamos nuestras teorías con la esperanza de eliminar las que no son verdad. De esta manera podemos conseguir nuestro propósito de perfeccionar nuestras teorías; incluso como instrumentos: haciendo redes que estén cada vez mejor adaptadas para capturar nuestro pescado, el mundo real. Sin embargo, nunca serán instrumentos perfectos para ese propósito. Son redes racionales hechas por nosotros mismos, y no deben confundirse con una representación completa del mundo real en todos sus aspectos, ni siquiera aunque tengan un gran éxito; ni siquiera aunque parezcan producir excelentes aproximaciones a la realidad [3].

---

[1] Véase especialmente Schlick, *Fragen der Ethik*, 1930, capítulo VII, sección 1, pp. 105 y ss.

[2] Ocurre que el primer ensayo que publiqué («Uber die Stellung des Lehrers zu Schule und Schüller», en *Schulreform*, 4, 1925, pp. 204-208) estaba dedicado, fundamentalmente, al mismo argumento que el que se desarrolla aquí a continuación.

[3] *Cf.* J. von Neumann y H. H. Goldstine, *op. cit.*, p. 1024: «Esto... está... estrechamente relacionado con la observación metodológica de que una formulación

Si tenemos bien presente que nuestras teorías son nuestra propia obra, que somos falibles y que nuestras teorías reflejan nuestra falibilidad, entonces dudaremos de que las características generales de nuestras teorías, tales como su simplicidad o su determinismo *prima facie,* correspondan a las características del mundo real.

Lo que quiero decir es esto: si hemos contrastado un enunciado como «todos los perros tienen rabo» y ha resistido nuestras contrastaciones, entonces quizá todos los perros (o gatos) tengan rabo, o al menos, aproximadamente todos. Pero sería un error concluir, del hecho de que una oración de sujeto-predicado tan universal haya tenido bastante éxito en su descripción del mundo, o incluso del hecho de que sea verdadera, que el mundo tiene una estructura de sujeto-predicado, o que está formado por sustancias que tienen ciertas propiedades. De la misma manera, el éxito o incluso la verdad, de enunciados muy simples, o de enunciados matemáticos, o de enunciados en inglés, no debe tentarnos a inferir que el mundo es intrínsecamente simple, o matemático, o británico [4]. Todas esas inferencias han sido extraídas, de hecho, por algún filósofo que otro; pero, reflexionando sobre ellas, tienen poco de recomendables. El mundo, tal como lo conocemos, es enormemente complejo; y, aunque pueda tener aspectos estructurales que sean simples en un sentido u otro, la simplicidad de algunas de nuestras teorías —que son nuestra propia creación— no entraña la simplicidad intrínseca del mundo.

La situación respecto al determinismo es similar. La teoría de Newton, consistente en la ley de la inercia, la ley de la gravedad, etcétera, puede ser verdad, o aproximarse mucho a la verdad; es decir, el mundo puede ser como afirma la teoría que es. Pero no hay un enunciado de determinismo en esta teoría; la teoría no afirma en ningún momento que el mundo esté determinado; antes bien, es la *propia teoría* la que tiene ese carácter que yo he llamado «determinista *prima facie*».

Ahora bien, el carácter determinista *prima facie* de una teoría está estrechamente ligado a su simplicidad; las teorías deterministas *prima facie* son relativamente fáciles de contrastar y las contrastaciones pueden ir haciéndose progresivamente más precisas y severas. De mis consideraciones sobre el contenido, la contrasta-

---

matemática representa necesariamente sólo una teoría (más o menos explícita) de alguna fase (o aspecto) de la realidad, y no la realidad misma.»

[4] *Cf.* mi artículo «¿Qué es la dialéctica?», *Mind,* 49, N. S., p. 414, esp. p. 420, reimpreso en *El desarrollo del conocimiento científico [G. & R.].*

bilidad y la simplicidad, se sigue que las teorías de este tipo deberían preferirse a otras; que es por lo que tratamos de construir éstas preferentemente y perseveramos en ellas (si resisten las contrastaciones) siempre que los problemas que tenemos planteados nos lo permitan [5].

Al mismo tiempo, no parece más justificable inferir de su éxito que el mundo tiene un carácter intrínsecamente determinado que inferir que el mundo es intrínsecamente simple.

El método de la ciencia depende de nuestros intentos de describir el mundo con teorías simples: las teorías complejas pueden resultar incontrastables, aunque sean verdaderas. La ciencia puede describir como el arte de la ultra-simplificación sistemática el arte de discernir lo que puede ser ventajoso omitir.

Es importante ver la conexión entre este resultado y el problema de la capacidad de dar razón. Consideremos una tarea de predicción tal como el cálculo del estado de nuestro sistema solar dentro de tres meses, con una precisión estipulada de las posiciones finales y de los momentos finales de los planetas. Si, en el sentido de la capacidad de dar razón, queremos calcular la imprecisión permisible de las condiciones iniciales, entonces necesitamos no sólo la mecánica de Newton, sino también un *modelo* de nuestro sistema solar. En otras palabras, necesitamos una lista de los planetas, de sus masas, posiciones y velocidades; es decir, necesitamos una descripción aproximada del estado del sistema hoy. Pero al dar esa descripción, tenemos que hacer uso, invariablemente, de nuestra teoría.

Primero, es la teoría la que decide qué es lo que pertenece al estado del sistema (posiciones, masas, velocidades) y qué es lo que no pertenece a él (por ejemplo, los diámetros de los planetas; sus temperaturas; sus capacidades térmicas, y sus propiedades químicas y magnéticas).

Segundo, la teoría nos dice qué tamaños de «planetas» pueden descuidarse (los meteoritos, por ejemplo). Dicho de otro modo, es absolutamente ingenuo pensar de «toda tarea de predicción» como si se tratara de todo estado concebible del mundo, o suceso en el mundo: toda tarea de predicción, y especialmente toda tarea de predicción que pueda ser concebiblemente capaz de dar razón, opera con un *modelo* simplificador. Esta es la consecuencia de ver el mundo a la luz de una teoría simplificadora; y cualquier cosa

[5] *Cf. La lógica de la investigación científica*, final de la sección 36.

que no sea iluminada por este faro permanece oscura: es descuidada [6].

La universalidad de nuestras teorías plantea problemas similares. Tenemos muchas razones para creer que el mundo es único: una única y enormemente compleja —quizá incluso infinitamente compleja— combinación de ocurrencias de procesos en interacción. Sin embargo, tratamos de describir este mundo sin par con la ayuda de teorías *universales*. ¿Describen esas teorías las características universales del mundo, las regularidades?, o ¿es la universalidad, puramente, como la simplicidad, una característica de nuestras teorías —quizá de nuestro lenguaje teórico—, pero no del mundo?

Yo creo que el caso es aquí algo diferente del de la simplicidad. Si decimos «todos los perros tienen rabo», entonces de hecho estamos afirmando algo sobre todos los perros; esto está claro porque tendremos que retractarnos de ese enunciado si encontramos una raza de perros sin cola (análoga a la raza de los gatos sin cola de la Isla de Man). Así pues, la universalidad es algo que afirman nuestras teorías y algo que nosotros nos proponemos contrastar. La simplicidad, por el contrario, no es afirmada por nuestra teoría; y, si lo fuera, podríamos no saber cómo contrastarla [7].

Al mismo tiempo, es sólo nuestro intento de *explicar* el mundo, es decir, de describirlo en términos de teorías cada vez más universales, lo que nos hace subir en la escala, no sólo de niveles de universalidad, sino también de niveles de aproximación. Las teorías que pueden explicárnoslo con la ayuda de teorías de mayor universalidad parecen a menudo, desde el nuevo nivel, sólo aproximaciones.

Es concebible que este proceso de aproximación pueda llegar a su fin un día, porque puede ser que un día lleguemos a *la* teoría

---

[6] [*(Añadido en 1981.) He hablado en varios sitios de la teoría de la ciencia como faro (*La sociedad abierta —O. S.*, 1945—, vol. II; *Conocimiento objetivo —Obj. Kn.*, 1972—, Apéndice I; *El desarrollo del conocimiento científico —C. & R.*, 1963—. Más recientemente, he hablado también de la «teoría del lenguaje humano como faro» en relación con el problema formulado en *El desarrollo del conocimiento científico (C. & R.)*. El sentido de la metáfora es que el faro *puede* iluminar algún aspecto del mundo (incluso desde una posición escogida) sin influirlo ni interferir con él en modo alguno: la objetividad de la verdad de un enunciado puede ser absoluta, incluso aunque el hecho que describa no pueda ser abstraído (o sacado) de la realidad excepto con la ayuda del lenguaje descriptivo humano.]

[7] *Cf.* también *La lógica de la investigación científica [L. Sc. D.]*, sección 15 y Apéndice *vii.

verdadera y completa del mundo (aunque, en el momento actual, ese día parece estar muy lejos —mucho más de lo que parecía estar en tiempo de Kant o de Laplace). Pero, incluso aunque encontrásemos la teoría verdadera del mundo, no podríamos saber de ninguna manera —como lo comprendió Jenófanes [8]— que la habíamos encontrado. Durante casi dos siglos, la teoría de Newton se consideró *la* verdadera teoría del mundo; e, incluso si encontramos una teoría que nos parezca tan satisfactoria como pareció la de Newton a la mayor parte de los físicos durante esos dos siglos, no deberemos sentirnos nunca seguros de que no pueda encontrársele algún día alguna objeción seria.

Así, no deberíamos descartar la posibilidad de que tengamos que conformarnos con seguir perfeccionando para siempre nuestras aproximaciones. Puede que sea consecuencia de *la unicidad del mundo* que el intento de describirlo en términos de teorías universales nos lleve a una secuencia infinita de aproximaciones, un intento parecido al de describir un número irracional en términos de razones de números naturales [9]. Nuestro intento de describir el mundo en términos de teorías universales puede ser un intento de racionalizar lo único, lo irracional, en términos de leyes universales hechas por nosotros mismos [10]. (Difiere del método de aproximación por una secuencia de razones en que cada paso de aproximación parece describir un *aspecto parcial* del mundo sin el cual no podríamos entender el próximo.) Nuestro propio método de ultra-simplificación puede dar lugar a la brecha que tratamos de cerrar con nuestras aproximaciones. Pero, puesto que no hay medida absoluta para el grado de aproximación logra-

---

[8] Véase Diels-Krantz, 21 [11], fragmento 34 (6.ª edición, 1951, p. 137). [Véase también Popper, «Retorno a los presocráticos», en *El desarrollo del conocimiento científico (C. & R.,* 1963). *Ed.*]

[9] Es interesante que J. B. S. Haldane, partiendo esencialmente de la premisa de que nuestras leyes (u otras «leyes que puedan, en principio, formularse matemáticamente») no son solamente aproximaciones sino que son estrictamente verdaderas, llega a la conclusión de que nuestro mundo no es único, sino repetitivo, de modo que, por ejemplo, «cada dos personas que se encuentran en el mundo en la vida actual... lo harán, por tanto, un número infinito de veces». (Véase «Consequences of Materialism», en *The Inequality of Man,* Pelican Books, 1937, pp. 169 y 168.) Esas especulaciones son, a los propios ojos de Haldane, «extrañas» (p. 170); y puede apelarse a su carácter extrañamente metafísico para apoyar la metafísica contraria y, creo yo, menos fantástica: la teoría de la unicidad del mundo y del carácter aproximado de todas las teorías.

[10] Con la concepción de que el mundo es irracional y de que es tarea de la ciencia racionalizarlo, compárese la nota 19 del capítulo 24 de mi *Sociedad abierta [O. S.].*

do por la tosquedad o la fineza de nuestra red —sino sólo una comparación con aproximaciones mejores o peores—, incluso nuestros esfuerzos de mayor éxito no pueden producir más que una red cuya malla es demasiado tosca para el determinismo. Intentemos examinar el mundo exhaustivamente con nuestras redes; pero sus mallas siempre dejarán escapar algún pequeño pez: siempre habrá suficiente juego para el indeterminismo.

Esto puede verse con mayor claridad en lo que más nos concierne. No hay el menor signo de que, con la ayuda de los métodos científicos, podamos acercarnos realmente a una descripción científica o a una clasificación de las personalidades humanas: continúan siendo únicas, a pesar de todas las tentativas de clasificación y medición.

16. COMPARACION CON LA PERSPECTIVA DE KANT

La mayoría de las consideraciones que se han avanzado en la sección precedente son de carácter lógico o metodológico. Pero la concepción del mundo puede muy bien describirse como metafísica: corresponde estrechamente a la idea del mundo de los *noumenos* o de *las cosas en sí mismas.*

Kant no sólo creía, como la mayoría de sus contemporáneos —astrónomos y físicos incluidos—, en la verdad de la teoría de Newton; incluso creía que ésta era válida *a priori* [11]. La posibilidad de que la teoría de Newton fuese sólo una espléndida aproximación no se les ocurrió, y yo creo que no podía ocurrírseles. Esta es la razón por la que él distinguió el mundo de la apariencia o *«Naturaleza»* —el mundo sobre el que nuestro intelecto impone *a priori* sus leyes (newtonianas)— del mundo de las cosas en sí mismas, el mundo real de los *noumenos.* La Naturaleza, el mun-

[11] Véase especialmente Kant, *Metaphysical Foundations of Nat. Science,* 1786 (publicado el año anterior a la segunda edición de la *Crítica de la razón pura).* Es bien sabido, naturalmente, que Kant era newtoniano, lo que no es tan bien sabido —aunque es de importancia crucial para la comprensión de su filosofía— es que cuando hablaba de «la ciencia natural pura» se refería a la dinámica de Newton, a la que consideraba válida *a priori* (véase mi artículo sobre «La naturaleza de los problemas filosóficos y sus raíces en la ciencia», *The British Journal for the Philosophy of Science,* 3, 1952, esp. pp. 153 y ss., e «Immanuel Kant, Philosopher of the Enlightenment», *The Listener,* 18 de febrero de 1954). [Véase *El desarrollo del conocimiento científico (C. & R.),* capítulos 2, 7 y 8. *Ed.*]

do en el tiempo y en el espacio, estaba sujeta a leyes causales, creía él; leyes que lo determinaban todo en la naturaleza «con necesidad». Nuestras acciones en el espacio y en el tiempo estaban completamente predeterminadas; podrían calcularse de antemano, como los eclipses. Sólo como *noumenos,* como cosas en nosotros mismos, podíamos ser libres.

Si sustituimos el mundo de los *noumenos* de Kant por nuestro mundo de cosas o procesos considerados bajo el aspecto de su unicidad, y el mundo de los fenómenos de Kant por nuestro mundo de cosas consideradas bajo el aspecto de la universalidad, nos acercamos a la opinión desarrollada en la sección precedente; excepto que, como he mostrado, hay que distinguir entre causalidad y determinismo, y que nuestro mundo de unicidad es —al contrario que el mundo noumenal de Kant— en el espacio e, incluso más importante, en el tiempo; porque yo considero de una importancia crucial el distinguir entre el *pasado* determinado y el *futuro* abierto.

Así, yo estoy de acuerdo con Kant cuando indica que una teoría como la de Newton es una creación nuestra, impuesta, como dice él, por nuestro intelecto sobre la naturaleza; que, de esta manera, nuestro intelecto racionaliza la naturaleza, y que hay una realidad —más profunda que la descrita por la teoría de Newton o por cualquier otra teoría— a la que no debemos considerar determinista. Pero no estoy de acuerdo con la creencia de Kant de que la teoría de Newton tiene que ser verdadera, y de que una teoría que nosotros imponemos a la naturaleza tiene, por esa razón, que ser válida *a priori* o de un carácter determinista *prima facie.* Tampoco estoy de acuerdo con su opinión de que la propia realidad indeterminista no puede ser conocida. Aunque el mundo único en el que vivimos no puede ser nunca conocido por completo, nuestro conocimiento científico es un intento —y un intento de un éxito sorprendente— de conocerlo cada vez mejor. En este sentido del conocimiento, todo nuestro conocimiento se refiere sólo a este único mundo nuestro y a ningún otro.

La dificultad fundamental de la solución de Kant —que como cosas libres en sí mismas nosotros no estamos en el espacio ni en el tiempo, mientras que nuestras acciones están en el espacio y en el tiempo, y, por tanto, determinadas— es claro que no se plantea en mi solución. Así, resulta posible decir que estamos tomando decisiones morales aquí y ahora (lo que no dudo que Kant hubiera deseado poder decir).

Kant expresa su determinismo en este pasaje [12]:

> Así pues, podemos admitir la verdad de la concepción según la cual podríamos calcular de antemano y con certeza —como lo hacemos con los eclipses lunares o solares— el comportamiento futuro de cualquier hombre, si tuviéramos un discernimiento tan profundo de sus modos de pensar como para conocer todas sus fuentes de acción más íntimas y también todas las circunstancias externas pertinentes: y, sin embargo, al mismo tiempo, podemos afirmar que el hombre es libre.

Este pasaje es testimonio de la fuerza de la creencia de Kant en el indeterminismo: incluso era más fuerte que su creencia errónea de que la ciencia (la ciencia *a priori)* nos obligaba a aceptar el determinismo. Porque lo que él dice aquí de la predictibilidad es claramente puro determinismo, como él mismo pone de relieve. Naturalmente, su fórmula puede ser salvada muy sencillamente por la observación de que nunca podemos tener «un discernimiento tan profundo» como el que sería necesario para la tarea de predicción. Pero aunque esto salvaría su fórmula, satisfaciéndola vacuamente, no salvaría lo que él quería decir; además, eso supondría abandonar el principio de poder dar razón y, con él, el determinismo «científico».

17. ¿ES LA FISICA CLASICA CAPAZ DE DAR RAZON?

Los argumentos filosóficos desarrollados en la sección 15 y comparados con los de Kant en la precedente sugieren algunas consecuencias de una naturaleza ligeramente más técnica: sugieren un modo de mostrar que la física clásica no es capaz de dar razón, incluso con independencia de los resultados, decisivos, de Hadamard.

La importancia de esos resultados está críticamente limitada. No tienen por qué afectar a la visión determinista, ni siquiera a la visión mecanicista del mundo, tan cara a un newtoniano. Pueden ser perfectamente válidos y, sin embargo, no sorprender ni disgustar a un newtoniano. Pero sí que afectan al determinismo «científico»; es decir, a la concepción de que el determinismo está respaldado por la ciencia humana, por la experiencia humana;

---

[12] *Critique of Practical Reason*, ediciones 4.ª a 6.ª, p. 172, ed. Cassirer, vol. 5, página 108. Véase también la nota 13 de la sección 7, más arriba.

porque esta forma de determinismo está decididamente ligada al principio de poder dar razón.

Para que una tarea de predicción sea capaz de dar razón, tenemos que tener un *modelo* del sistema (como indiqué en la sección 15); es decir, *una descripción aproximada de su estado*. Esto es claro si consideramos que para resolver un problema de-un-cuerpo o uno de-dos-cuerpos, o digamos que un problema de-tres-cuerpos en el que la interacción de dos de los tres cuerpos es despreciable en una primera aproximación (a causa de su gran distancia y de su poca masa), no necesitaremos unas condiciones iniciales tan precisas como las que necesitamos para resolver, con el mismo grado de precisión, un problema de-tres-cuerpos con una potente interacción entre dos *cualesquiera* de los tres cuerpos. Sin embargo, si necesitamos que nos den el estado inicial aproximado del sistema antes de que podamos empezar siquiera a calcular el grado de aproximación requerido por el principio de poder dar razón, entonces todo el problema de poder dar razón puede convertirse en indeterminado, para algunos casos, si no en insoluble. Porque surge la cuestión: ¿cómo tiene que ser el modelo de bueno para que nos permita calcular la aproximación que requiere el principio de poder dar razón? Puesto que la bondad del modelo es su grado de aproximación o precisión, estamos amenazados por una retrogresión infinita; y la amenaza será muy seria para los sistemas que sean complejos. Pero la complejidad del sistema tampoco puede evaluarse más que si se tiene a mano un modelo aproximado; una consideración que, de nuevo, indica que estamos amenazados por una retrogresión infinita.

Sin duda, en muchos casos que no son demasiado complejos, será posible proceder por el método siguiente: primero obtenemos un modelo, que puede ser bueno o malo: no necesitamos saberlo. Luego tratamos de calcular, según el principio de poder dar razón, la precisión que se requiere de las condiciones iniciales necesarias para completar nuestra tarea de predicción; y, si fracasamos porque el primer modelo que teníamos no era bueno, tratamos de conseguir un modelo mejor.

Este método puede tener éxito con frecuencia; y, si lo tiene, diremos, ciertamente, que el principio de poder dar razón está satisfecho. Pero ¿qué pasa si volvemos a fracasar, incluso con un modelo mejor? Está claro que debemos limitar, de antemano, o bien el número permisible de exigencias que debe cumplir un modelo más perfecto, o bien la «bondad», es decir, el grado de preci-

sión que podemos pedir al modelo. Pero la tarea de calcular cualquiera de estas cosas nos puede llevar simplemente a un problema de capacidad de dar razón de un orden superior. Y, con ello, podemos muy bien encontrarnos camino de una retrogresión infinita. Porque no hay razón alguna para creer que el problema de orden superior sea más fácil de resolver que el problema de orden inferior, o que se necesite un modelo menos bueno para su solución que para la solución del problema de orden inferior. Tampoco hay ninguna razón para creer que los métodos de aproximación vayan a seguir siendo capaces de perfeccionar sus resultados indefinidamente.

No se ofrecen estas consideraciones como decisivas, sino más bien como indicaciones de que, en relación con el problema del determinismo «científico», el problema de la complejidad puede afectar de modo crucial a la situación y de que la complejidad del mundo real puede destruir todos los argumentos que mantienen que el determinismo se basa en la experiencia científica o en el éxito de nuestras teorías científicas.

Aparte de esas consideraciones tan generales, hay otras más específicas, en relación con la capacidad de dar razón en los dos sentidos, el más fuerte y el más débil (tal como se definieron en la sección 3).

Primero, respecto a una débil capacidad de dar razón, ocurre que, incluso aunque se nos den unas condiciones iniciales exactas, sólo en casos especiales podemos predecir el futuro de un sistema newtoniano que conste de más de dos cuerpos, y no parece haber esperanza alguna de resolver esta tarea para más de tres cuerpos, a menos que el sistema pertenezca a esas estructuras, muy especiales, a las que son aplicables ciertos métodos de aproximación. No sabemos cómo abordar un sistema que conste de, digamos, ocho, u ochenta, u ochocientos cuerpos aproximadamente iguales, situados a distancias aproximadamente iguales. Puesto que, en el momento actual, no tenemos método alguno para calcular predicciones para un sistema de ese tipo de complejidad, no tenemos, *a fortiori*, método de descubrir cómo tendría que ser de preciso cualquier conjunto dado de condiciones iniciales para poder resolver una tarea de predicción con un grado predeterminado de precisión.

Mientras no haya perspectivas serias de resolver el problema general de-*n*-cuerpos de la dinámica de Newton, no hay ninguna razón para creer que la dinámica de Newton es capaz de dar razón, ni siquiera en el sentido más débil de «capaz de dar razón».

Aún más, si procedemos a examinar la situación con respecto a la capacidad de dar razón en su sentido más fuerte, podemos encontrar buenas razones para creer que la dinámica de Newton *no* es capaz de dar razón.

Consideremos un sistema gravitatorio newtoniano (aproximadamente) aislado, muy alejado en el espacio vacío, consistente en una serie de cuerpos más bien pequeños (de masas, digamos, entre unas cuantas toneladas, los más pequeños, y unas cuantas docenas de toneladas, los mayores). Y consideremos cómo podríamos determinar, por medio de mediciones, las condiciones iniciales que se necesitan para predecir un sistema de este tipo, y más especialmente las masas de los diversos cuerpos pertenecientes a él. No podemos hacer uso de un péndulo ni de un cuerpo de contraste combinado con una balanza de resorte, porque no podemos introducirnos en un sistema de ese tipo sin perturbarlo seriamente y de una forma impredictible. (La perturbación es impredictible porque sabemos demasiado poco acerca del sistema que, debido a nuestra intrusión, puede desequilibrarse tanto que algunos de sus miembros pueden escaparse de él antes de que podamos medirlos.) Así pues, debemos suponer que podemos hallar las condiciones iniciales de un sistema de este tipo observándolo *visualmente* desde fuera, como un sistema estelar. Podemos suponer que el sistema o bien tiene sus propias fuentes de luz visible o es iluminado por luz de fuera; y podemos suponer que, usando luz visible, no perturbamos el sistema. (Este supuesto es razonable, puesto que los cuerpos del sistema son cuerpos macroscópicos, con suficiente peso como para no ser perturbados en forma apreciable por mediciones con luz visible; en contraposición con la concepción habitual de Heisenberg en relación con los átomos o las partículas sub-atómicas.) Podemos suponer incluso que nos es posible medir ópticamente las velocidades (con ayuda del efecto Doppler) en lugar de medir las posiciones y un intervalo de tiempo. (Por ejemplo, podemos observar ópticamente tres vectores de velocidad desde tres planetas distantes no-coplanarios entre los que se haya establecido comunicación.) Para calcular las masas, o, *al menos*, las razones-masa, tenemos que usar la ley inversa del cuadrado y medir (digamos que por radar) para el mismo instante de tiempo distancias y aceleraciones [13].

[13] Véase C. G. Pendse, *Phil. Mag.*, 7.ª serie, 24, pp. 1012 y ss., 1937; 27, pp. 51 ss., 1939; H y 29, pp. 477 y ss., 1940. Pendse investiga la teoría de la determina-

Consideremos ahora cómo podemos medir las *aceleraciones* ópticamente. La única manera consiste en medir velocidades y ver cómo cambian. Pero hay un problema, incluso para medir velocidades instantáneas: con cuanta más precisión queramos determinar las velocidades, menos precisa será la determinación del instante al que pertenecen. Pero incluso si (como concesión a nuestros oponentes) estuviéramos dispuestos a despreciar esa dificultad, aparece de nuevo en una forma más severa si medimos la aceleración. Porque, para hacerlo, tenemos que medir las velocidades en dos instantes de tiempo separados por un intervalo de tiempo finito, no demasiado corto; si no, no seremos capaces de observar diferencia apreciable y, en consecuencia, no mediremos la aceleración; no obstante, si tomamos con intervalo no demasiado corto, entonces no podemos atribuir las aceleraciones a ningún instante preciso, y obtenemos, por añadidura, simples aceleraciones medias.

La expresión matemática de esto puede simplificarse como sigue:

Para la luz proveniente de una fuente, tenemos la fórmula general

$$\Delta\nu\Delta t = 1 \tag{1}$$

Si aplicamos esto al efecto Doppler, que puede expresarse

$$\nu = \lambda_0 (\nu_0 - \nu_1) \tag{2}$$

entonces encontramos que podemos determinar ν_0 y, por tanto, λ_0 con tanta precisión como deseamos: se refieren a la luz de una fuente constante que podemos observar, en principio, durante todo el tiempo que queramos. Así, podemos dar a ν_0 y λ_0 como conocidas con exactitud. Pero el cambio $\nu_0 - \nu_1$ no puede medirse con más precisión que ν_1. Así, combinando (1) y (2), obtenemos

$$\Delta\nu\Delta t = \lambda_0 \tag{3}$$

---

ción de las razones-masa por medio de aceleraciones. En sus comentarios sobre la obra de Pendse, V. V. Narlikar (*Phil. Mag.*, 27, pp. 33 y ss.) señala correctamente que, en presencia de la ley inversa del cuadrado de la fuerza, las razones-masa pueden obtenerse a partir de las aceleraciones (un punto que no había sido mencionado previamente por Pendse, y con el cual estuvo de acuerdo en el último de esos artículos). Ninguno de estos autores considera la dificultad inherente a medir las aceleraciones no-constantes de Newton en instantes definidos de tiempo.

Como hemos supuesto que operamos con luz visible (este supuesto puede debilitarse algo, pero hay un límite inferior para λ_0 porque una luz demasiado fuerte tendría una potencia de penetración demasiado grande), habrá un límite inferior de λ_0; y, en consecuencia, vemos por (3) que no podemos hacer independientemente a Δv y Δt tan pequeños como queramos. Sin embargo, (3) sólo es válida para velocidades constantes; para velocidades variables, la situación con respecto a v es incluso peor que la expresada por (3). Porque si v cambia, como debemos suponer que hace, puesto que estamos interesados en medir la aceleración, hay dos fuentes diferentes, y cada una tiende a incrementar el valor de Δv; primero, la elección de un Δt *pequeño*, es decir, de un corto período de medición (o de exposición de la placa fotográfica en la que registramos el cambio), que incrementa Δv de acuerdo con (3) o (1); segundo, la elección de un Δt *grande*, con la consecuencia de que v y v_1 cambien durante el período de medición, de forma que v_1 queda emborronada y, por tanto, Δv, grande. Habrá, por consiguiente (en el mejor de los casos), un valor óptimo de Δt, dependiendo de la aceleración, tal que si Δt es suficientemente grande [desde el punto de vista de la fórmula (3)] y al mismo tiempo no demasiado grande, como para permitir que la aceleración incremente Δt. Correspondiendo a ese valor óptimo de Δt, habrá un valor mínimo de Δv, un valor que no podemos disminuir.

Si intentamos ahora determinar la aceleración a basada en una segunda medición (de v_2 hecha en el tiempo t_2), de forma que

$$a \approx \frac{v_2 - v_1}{t_2 - t_1} \tag{4}$$

entonces vemos en seguida que, si el intervalo $t_2 - t_1$ no es considerablemente mayor que Δt, obtenemos un valor completamente indeterminado, tal como $a = 0/0$. Así, resulta imposible, en principio, determinar a con precisión para cualquier intervalo corto de tiempo Δt; y sólo podemos obtener un valor medio de a para un período mucho más largo: $t_2 - t_1$, donde t_1 y t_2 están determinados simplemente con una precisión limitada Δt, y donde incluso el valor *medio* de a no está determinado más que de forma imprecisa, debido a la impresión irreductible de v_1 y v_2 ($v_2 - v_1$ tiene también que ser mucho mayor que Δv).

Queda claro tras estas consideraciones que no podemos, con ayuda de luz visible, medir en nuestro sistema newtoniano, con toda la precisión que deseemos, todas las diversas aceleraciones

en un cierto instante (a determinar con tanta precisión como queramos). En consecuencia, no podemos determinar las razones-masa de los cuerpos con tanta precisión como queramos. No parece, por tanto, que ni siquiera en todos los sistemas clásicos macroscópicos sean posibles las mediciones que podrían darnos unas condiciones iniciales tan precisas como las quisiéramos; y esto lleva, una vez más, a la conclusión de que no todas las tareas de predicción de la física clásica pueden realizarse sobre la base de las mediciones de las condiciones iniciales.

A fortiori, esto significa, además, que la física clásica no es capaz de dar razón en el sentido más fuerte de «capaz de dar razón».

La similaridad entre la situación en la física clásica aquí descrita y el principio de indeterminación que, según Heisenberg, es válido para la teoría cuántica, es demasiado obvia para requerir comentario. [Las fórmulas de indeterminación de Heisenberg son, por supuesto, iguales que la fórmula (1): resultan, simplemente de la multiplicación de los dos miembros de (1) por h o por $\hbar$.] La similaridad es consecuencia de aplicar, como Heisenberg, un análisis operacionalista. Pero, como, al contrario que Heisenberg, yo no soy operacionalista, no uso esas consideraciones para derivar consecuencias ontológicas; por el contrario, sólo quiero señalar las dificultades inherentes a la concepción de que la mecánica de Newton difiere de la teoría cuántica en que es determinista.

18. EL PASADO Y EL FUTURO

El carácter aproximado de todo conocimiento científico —la red cuya malla tratamos de hacer cada vez más fina— proporciona lo que me parece a mí que es el argumento filosóficamente más fundamental en contra del determinismo «científico» y en favor del indeterminismo. Después de él, pero todavía de importancia, está *el argumento de la asimetría entre el pasado y el futuro.*

Uno no puede cambiar el pasado, aunque se han hecho intentos de hacer lo más parecido a ello (una cosa que, según el idealismo o el subjetivismo o el positivismo sería lo mismo): cambiar nuestro conocimiento del pasado mediante el falseamiento de los documentos históricos existentes. Puesto que el pasado no es

más que lo que ha ocurrido, es una verdad trivial decir que *el pasado está completamente determinado por lo que ha ocurrido.* La doctrina determinista —según la cual el porvenir está completamente determinado por lo que ha ocurrido— destruye caprichosamente una asimetría fundamental en la estructura de nuestra experiencia [14], y está en absoluto conflicto con el sentido común. Todas nuestras vidas, todas nuestras actividades, están dedicadas a intentar afectar el futuro. Está claro que creemos que lo que ocurrirá en el futuro está en gran manera determinado por el pasado y el presente, ya que todas nuestras acciones racionales actuales son intentos de influir, o determinar, el futuro. (Esto es cierto incluso en los intentos de falsear el pasado.) Pero está igual de claro que consideramos que el futuro no está todavía completamente fijado; al contrario que el pasado, que está cerrado, por decirlo así,

[14] La asimetría a la que aludo entraña, claramente, la *«flecha del tiempo»*, pero entraña más que la afirmación de que el tiempo tiene una dirección (una «flecha»). Porque esta flecha se ha encontrado en el pasado y se encontrará en el futuro: su existencia no es suficiente para establecer la diferencia fundamental entre el pasado y el futuro que yo tengo en mente. Respecto a *la flecha del tiempo*, es, en mi opinión, un error hacer a la segunda ley de la termodinámica responsable de su dirección. Incluso en un proceso no termodinámico, tal como la propagación de una onda desde un centro, es, de hecho, irreversible. Supongamos que se toma una película de una amplia superficie de agua inicialmente en reposo a la que se ha tirado una piedra: ningún físico confundirá el final de la película con el principio; porque la creación de una onda circular, que va contrayéndose seguida de una zona de agua tranquila acercándose al centro, sería (considerada causalmente) milagrosa. Si se contesta a esto diciendo que el fenómeno supuestamente milagroso podía haber sido causado por una multiplicidad de osciladores colocados en un inmenso círculo y funcionando al unisono («coherentemente»), entonces yo respondería que esta disposición sería milagrosa a menos que estuviese dirigida o coordinada *desde un centro*, por ejemplo, por medio de señales luminosas, de modo que, para establecer la posibilidad física y causal de la película invertida, tendríamos que suponer que está *precedida de un proceso en el cual las ondas no se contraen sino que se extienden*, como en la película original; un proceso que, si vuelve a invertirse la película, se convierte a su vez en milagroso, y crea así una retrogresión infinita. Dicho de otro modo, todas las causas (aunque no todas las fuerzas) se extienden desde centros (reminiscentes del principio de Huygens); y la flecha del tiempo se debe, en el caso que hemos considerado, no a las ecuaciones, sino al carácter de las condiciones iniciales. Véase también A. Einstein, *Physikalische Zeitschrift*, 10, 1909, página 821; y mis propias cartas, «The Arrow of Time», *Nature*, 177, 1956, p. 538, «Irreversibility and Mechanics», *Nature*, 178, 1956, pp. 381 y ss., e «Irreversible Processes in Physical Theory», *Nature*, 179, 1957, pp. 1296 y ss. [Véase también Popper, «Irreversibility; or Enthropy since 1905», *British Journal for the Philosophy of Science*, 8, 1957, pp. 151-155; «Erratum», *ibid.*, 8, 1957, p. 258; «Irreversible Processes in Physical Theory», *Nature*, 181, 1958, pp. 402-3; «Time's Arrow and Enthropy», *Nature* 207, 1965, pp. 233-4; «Time's Arrow and Feeding on Negentropy», *Nature*, 213, 1967, p. 320; «Structural Information and the Arrow of Time», *Nature*, 214, 1967, p. 322; «Autobiography of Karl Popper», en P. A. Schilpp, ed.: *The Philosophy of Karl Popper*, vol. I, pp. 124-129, y vol. II, páginas 1140-1144; y *Búsqueda sin término* [*U. Q.*, 1976, pp. 209-219. *Ed.*]

el futuro está todavía abierto a influencias; todavía no está completamente determinado.

Estoy lejos de afirmar que, en cuestiones de este tipo, el sentido común y las actitudes comunes sean los árbitros definitivos: si hay buenas razones, basadas en argumentos y especialmente en teorías científicas contrastables, para aceptar una opinión en conflicto con el sentido común, entonces no tengo la menor duda sobre qué actitud adoptar. Este, sin embargo, no es el caso aquí. Porque hay incluso buenas razones científicas, provenientes de una teoría determinista *prima facie* —la teoría especial de la relatividad— que apoyan la opinión de sentido común de la «apertura» del futuro.

19. EL VEREDICTO DE LA RELATIVIDAD ESPECIAL

Si tengo razón al afirmar la asimetría entre el pasado y el futuro —lo cerrado del pasado y lo abiero del porvenir—, entonces esta asimetría tiene que poder representarse en la estructura de la teoría física.

Esta exigencia está perfectamente satisfecha por la teoría especial de la relatividad de Einstein [15]. En esta teoría existe, para cada observador —o, como yo prefiero decir, para cada sistema de inercia local—, un pasado absoluto y un futuro absoluto (que están separados por toda una región de posible contemporaneidad). El pasado (absoluto) del sistema es la región formada por todos los puntos espacio-temporales desde los cuales las influencias físicas (por ejemplo, las señales de la luz) pueden afectar al sistema; su futuro (absoluto) es una región formada por todos los puntos sobre los cuales el sistema puede ejercer una influencia física. En la representación geométrica de Minkowski, este pasado y este futuro forman los dos conos (más exactamente las dos partes de un doble cono cuatridimensional); su vértice A es el «aquí-y-ahora».

---

[15] Desgraciadamente, la posición es diferente con respecto a la teoría general de la relatividad de Einstein: aquí la asimetría de la teoría especial pasa a ser un asunto local. Pero Einstein lo reconoció indirectamente como una imperfección que hay que suprimir (incluso *ad hoc*, si no se dispone de un método mejor). Véase *Albert Einstien, Philosopher-Scientist*, ed. P. A. Schilpp, 1949, esp. p. 687; es decir, la respuesta de Einstein a Kurt Gödel. Véase también la próxima nota. [Véase también *Búsqueda sin término (U. Q.)*, notas 201 y 202. *Ed.*]

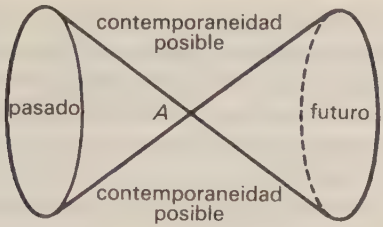

FIGURA 1

Un corte transversal del cono es así:

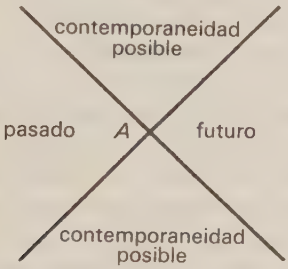

FIGURA 2

(He dispuesto este diagrama de forma que el tiempo apunte de izquierda a derecha, como es usual en las representaciones diagramáticas, aunque en los diagramas de relatividad es más usual hacer que el eje del tiempo apunte hacia arriba.)

No explicaré con detalle este conocido diagrama. Pero tengo que señalar que satisface plenamente la exigencia de asimetría entre el futuro y el pasado. En términos físicos, esta asimetría queda establecida por el hecho de que, desde cualquier lugar del «pasado», una cadena causal física (por ejemplo, una señal de luz) puede alcanzar cualquier lugar del «futuro»; pero desde ningún lugar del futuro puede ejercerse un efecto igual sobre ningún lugar del pasado [16].

[16] En este punto, la teoría general difiere de la especial, como ha mostrado Gödel *(cf.* la nota anterior). Sin duda, las soluciones mencionadas por Gödel no concuerdan con las ideas fundamentales de la teoría general y tendrían que excluirse como sugirió Einstein *(op. cit.,* p. 688). Pero una solución *ad hoc* no sería satisfactoria. Mientras no se obtenga la exclusión por medio de una modificación de las propias ecuaciones, puede dependerse del siguiente *«principio de la conexión inin-*

Como consecuencia de esto, el futuro se convierte en *«abierto»* para nosotros en el sentido de que no puede ser totalmente predicho *por nosotros,* mientras que el pasado está «cerrado»; es decir, la asimetría es del tipo que yo he tratado de establecer.

Para ver esto, supongamos que estamos en el vértice *A* y que queremos hacer una predicción completa sobre el estado de las cosas en nuestro sistema cuando *A* ha llegado al punto espacio-temporal *B*.

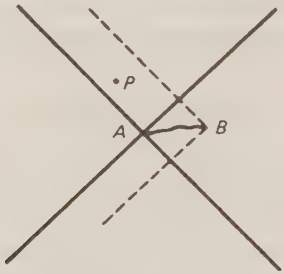

FIGURA 3

Es bien sabido [17] que no podemos hacer esto; como muestra la Figura 3, hay puntos como *P* que pertenecen al pasado de *B*, pero

terrumpida de las líneas-mundo» para lograrlo. Es un principio que me parece implícito en la propia idea de tiempo local al igual que en el principio de que la relatividad especial debe ser válida localmente dentro de la teoría general. En términos operacionales, mi principio puede formularse como sigue: *Cualquier «observador» (sistema material local) puede iniciar, en cualquier instante, un registro (pista causal); hacer anotaciones sucesivas en dicho registro, y disponer la conservación del registro durante cualquier período de tiempo finito deseado.* («Puede» significa lo que sigue: la posibilidad teórica de que cualquier línea-mundo que sea considerada coherente con las leyes de la naturaleza no tenga que entrañar la imposibilidad de las operaciones descritas en el principio mencionado más arriba.) Si se adopta este principio, entonces la existencia de un «observador» (un sistema material) cuya línea-mundo está cerrada en el tiempo lleva a una contradicción; porque, como puede mostrarse fácilmente, la existencia de una línea-mundo cerrada (que, por razones de coherencia, tiene que ser infinita y absolutamente repetitiva) entrañaría la destrucción periódica de cada uno de los registros (pista causal), puesto que, de otro modo, la pista (o curso) no sería totalmente repetitiva, sino que se enriquecería constantemente en cada renovación del itinerario cerrado. La adopción de mi principio haría que «el modo en que están dispuestos la materia y su movimiento en el mundo» *(op. cit.,* p. 562) dependiera de su estructura temporal (a cuya caracterización ayuda mi principio) en vez de ser al revés. El mismo resultado puede obtenerse —aún menos *ad hoc*— adoptando un «principio de indeterminismo»: esto también excluiría automáticamente todas las soluciones cosmológicas que permiten líneas-mundo cerradas.

[17] *Cf.* Herman Weyl, *Philosophy of Mathematics and Natural Science,* Princeton, 1949, pp. 102 y 210.

no al de *A;* lo que significa que desde *P* hay efectos que pueden alcanzar a *B;* pero es imposible para nosotros, en *A,* saber nada de las condiciones en *P,* ya que ningún efecto procedente de *P* puede alcanzarnos en *A; P* está fuera del cono-pasado de *A;* pero el cono-pasado de *A* es la única región sobre la que nosotros podemos tener conocimiento.

Quiero mostrar ahora que, como consecuencia de la asimetría entre el pasado y futuro, la relatividad especial no es ya determinista *prima facie* en el pleno sentido descrito más arriba. Lo haré mostrando que ya no hay un demonio laplaciano en la relatividad especial.

Tomemos de nuevo la situación de la Figura 3; *A* es nuestro presente y *B* es un punto espacio-temporal sobre el que se va a hacer una predicción. Los científicos humanos no pueden hacerla; pero suponemos que hay un demonio laplaciano, uno capaz de obtener *todas* las condiciones iniciales para una región del espacio suficientemente extensa (pero limitada) en un cierto instante de tiempo; es decir, para una cierta región de la que puede decirse que es «simultánea» en el sentido de la relatividad especial. En nuestra Figura 4, esta región está representada por el argumento de línea *C.*

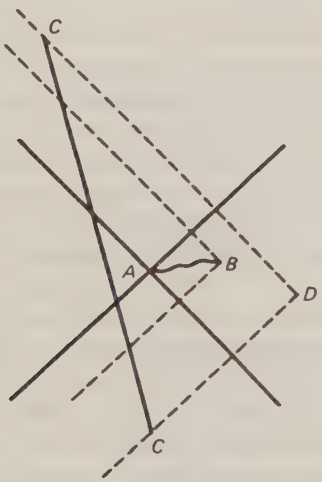

FIGURA 4

Está claro que para predecir el estado de las cosas en *B,* la línea *C* debe llegar *como mínimo* a las líneas punteadas que indi-

can el pasado de *B;* pero podemos suponer que va más allá de él. Así, *C* representa la región sobre la que el demonio ha recibido información completa. Ahora bien, dada esta región, la teoría nos permite hallar una posición espacio-temporal *D*, que, desde el punto de vista de la teoría, es la posición espacio-temporal *más temprana* en la que puede localizarse al demonio mientras recibe la información. Y *D* estará colocado de tal forma que *B* pertenezca al pasado de *D*. Esto significa que el demonio, cuando estaba calculando el estado de cosas en *B*, estaba haciendo una retrodicción en lugar de una predicción, en términos de la relatividad especial. Dicho de otro modo: *si tratamos de introducir al demonio laplaciano en la relatividad especial, encontramos que podemos calcular, a partir de la región de información del demonio, un límite inferior para la posición espacio-temporal* D *del demonio; y además encontramos que el demonio sólo calculó un suceso que estaba dentro de su propio pasado.*

Si la línea *C* se alarga infinitamente en las dos direcciones —lo que transforma a nuestro demonio limitado en un demonio ilimitado—, entonces vemos que el demonio puede, en efecto, calcular *cualquier* suceso. Pero eso ocurre porque él, en términos de la teoría, está situado en el futuro infinito, de forma que *cualquier* suceso pertenece a su pasado.

Así pues, el demonio de la relatividad especial no es ya el de Laplace; porque este demonio, al contrario que el de Laplace, no puede predecir, sino que sólo puede retrodecir.

En resumen, la relatividad especial convierte automáticamente a cada suceso sobre el que tenemos —nosotros o un demonio— alguna información definida en un suceso que pertenece a nuestro pasado o al pasado del demonio. De modo que puede decirse que, de acuerdo con la relatividad especial, el pasado es aquella región que puede, en principio, ser conocida; y que el futuro es aquella región que, aunque influida por el presente, está siempre «abierta»: no sólo es desconocida, sino que además, en principio, no es plenamente cognoscible, ya que el hacerse conocida por completo, incluso para un demonio, se convertiría en parte del pasado del demonio. La relatividad especial, a pesar de su carácter determinista *prima facie*, no puede, pues, usarse para apoyar al determinismo «científico» por dos razones: 1) Las predicciones requeridas por el determinismo «científico» deben interpretarse, desde el punto de vista de la propia relatividad especial, como *retrodicciones*. 2) Al ser retrodicciones, aparecen, desde el punto de vista

de la relatividad especial, como calculadas en el futuro del sistema predicho. Así, no puede decirse que han sido calculadas desde dentro de ese sistema: no satisfacen el principio de *predictibilidad desde dentro.*

La existencia de la teoría de la relatividad especial refuta, de este modo, el supuesto usual de que sería permisible inferir la verdad del determinismo «científico» de la verdad de una teoría determinista *prima facie.*

20. PREDICCION HISTORICA Y EL AUMENTO DEL CONOCIMIENTO

> No esperéis de mí, por tanto, ninguna profecía: si yo hubiera sabido lo que alguien iba a descubrir mañana, lo habría publicado hace mucho tiempo, para asegurarme la prioridad.
>
> HENRI POINCARÉ

Aparte de criticar el determinismo, he expuesto hasta ahora dos argumentos positivos en favor del indeterminismo: el del carácter aproximado del conocimiento científico y el de la asimetría entre el pasado y el futuro.

Ahora me ocuparé de un tercer argumento, menos fundamental quizá que cualquiera de esos dos, pero todavía muy importante, especialmente porque servirá para construir una refutación formal del determinismo «científico» (como se verá en la sección 23). Primero enunciaré el argumento en términos humanos. Muy sorprendentemente, puede volver a enunciarse por completo —y, de hecho, puede enunciarse de forma más precisa— desde un punto de vista puramente físico.

El núcleo del argumento es la consideración de que hay ciertas cosas sobre nosotros que no podemos predecir nosotros mismos por métodos científicos; más especialmente, *no podemos predecir, científicamente, los resultados que vamos a obtener en el transcurso del aumento de nuestro propio conocimiento.* Otros, que sean más sabios que nosotros, quizá puedan predecir el aumento de *nuestro* conocimiento, de la misma manera que nosotros podemos, en determinadas circunstancias, predecir el aumento del conocimiento de un niño; *pero tampoco ellos serán capaces de predecir o anticipar hoy lo que ellos mismos no sabrán hasta mañana.*

Esta formulación indica que puede haber una contradicción real implicada en la idea de predecir hoy lo que sólo mañana sabremos; y la hay. Pero no es fácil estar seguro de que esa contradicción no se debe puramente a nuestra formulación, y de que, de hecho, impide la posibilidad de una auto-predicción completa. Que lo hace, lo mostraré en las dos secciones siguientes.

Quiero señalar aquí algunas de las consecuencias del enunciado de que *no puede haber un científico capaz de predecir todos los resultados de todas sus propias predicciones.*

Una de esas consecuencias es que no será capaz de predecir algunos de sus estados futuros; y, aún más, no todos los estados de su propia «vecindad», es decir, de la parte de su entorno en la que él influye apreciablemente. Porque, si no sabe lo que sabrá mañana, tampoco puede saber cómo actuará mañana sobre su entorno. Así, el estado de su vecindad no será completamente predictible por él mismo *desde dentro,* aunque puede ser predictible *desde fuera,* por observadores que puedan predecir sus acciones, siempre que éstos no interfieran apreciablemente, con él ni con su vecindad.

De ahí se deduce que no hay sistema físico que pueda predecirse por completo desde dentro (y que la predictibilidad de sistemas del tipo del sistema solar debería describirse, según la terminología de la sección 11, como predictibilidad desde fuera).

El argumento puede usarse para *refutar la doctrina del historicismo,* la doctrina de que el cometido de las ciencias sociales es predecir el curso de la historia humana. Porque podemos argüir lo que sigue:

1. Si se puede mostrar que es imposible la auto-predicción completa, sea cual fuere la complejidad del predictor, entonces esto tiene que ser válido también para cualquier «sociedad» de predictores en mutua interacción; en consecuencia, ninguna sociedad de predictores en mutua interacción puede predecir sus propios estados futuros de conocimiento.

2. El curso de la historia humana está fuertemente influido por el aumento del conocimiento humano (la verdad de esta premisa tiene que ser admitida incluso por aquellos que, como los marxistas, ven en nuestras ideas, incluidas las ideas científicas, simples secuelas de ocurrencias materiales de uno u otro tipo).

3. No podemos, por tanto, predecir el curso futuro de la historia humana; no, en todo caso, de aquellos de sus aspectos que

estén influidos fuertemente por el aumento de nuestro conocimiento.

Este argumento no niega, naturalmente, la posibilidad de toda predicción social; por el contrario, es perfectamente compatible con la posibilidad de contrastar teorías sociales —por ejemplo, teorías económicas (pero no «teorías históricas»)— deduciendo de ellas predicciones que afirmen que ciertos sucesos tendrán lugar bajo ciertas condiciones y contrastando dichas predicciones [18].

21. PREDICCION DEL AUMENTO DEL CONOCIMIENTO TEORICO

Consideremos con mayor atención lo que una predicción del aumento del conocimiento científico puede implicar. Puede implicar nuestra capacidad de predecir ahora que en una fecha futura: a) *aceptaremos* (tentativamente, por supuesto) *como bien contrastadas ciertas teorías* que no son aceptadas actualmente y que quizá son, incluso, desconocidas en este momento; o b) deducir ciertas *explicaciones o predicciones* que son ahora desconocidas, a partir de teorías aceptadas ahora o que serán aceptadas más adelante, en conjunción con unas condiciones iniciales (que pueden ser todavía desconocidas).

En esta sección trataré de las predicciones del aumento del conocimiento teórico, es decir, del problema a). Aquí la cuestión más importante es si seríamos capaces de predecir la *aceptación de una teoría no aceptada anteriormente, sobre la base de nuevas contrastaciones.* De menor importancia es una cuestión preliminar: si seríamos capaces de predecir *el contenido de una teoría desconocida todavía,* la nueva idea que se le ocurrirá a alguien, o que será propuesta por alguien.

La razón por la que digo que esto es de menor importancia es que, a menudo, cuando una teoría es nueva, en el sentido de que

---

[18] Todos estos problemas se tratan, con algún detalle, en mi libro *La miseria del historicismo* [*Poverty,* 1957]; primero se publicaron en *Economica,* 1944-45, y en mi artículo «Predicción y profecía en las ciencias sociales», ahora recogido en el *Desarrollo del conocimiento científico [C. & R.],* capítulo 16. En *La miseria del historicismo,* sólo traté de mostrar que el historicismo es un método pobre; no traté de refutarlo. La refutación que se da aquí, en el texto, está implícita en mis artículos sobre «Indeterminismo», citados más arriba; está enunciada explícitamente en mi Prefacio a la edición francesa (París, 1956) de *La miseria,* y a la edición inglesa en forma de libro (1957).

ha sido aceptada recientemente, resulta que no es tan nueva como parece. Incluso puede haber sido propuesta hace mucho tiempo y olvidada porque no había evidencia a su favor, o no existía el problema para el que era necesaria, o el problema que ella podía resolver. Esto muestra que es importante, *si estamos interesados principalmente en al aumento del conocimiento «aceptado»,* en contraposición con el aumento de las nuevas ideas, es la aceptación tentativa de una teoría, sobre la base de nuevos problemas o nueva evidencia.

Consideremos, en primer lugar, la cuestión preliminar, del aumento de nuestras ideas teóricas. Un psicólogo, o, si se quiere, un fisiólogo, puede muy bien predecir las teorías o las expectaciones que un niño (o un animal), bajo la influencia de ciertos estímulos ambientales, puede formular y, después de ciertas contrastaciones, aceptar: el niño que se ha quemado (o el gato) teme al fuego. Si nuestro conocimiento psicológico (o fisiológico, o físico, o económico) es muy grande, podemos imaginar que somos capaces de aplicarnos un método similar a nosotros mismos, y predecir hoy las teorías que se nos van a ocurrir *por primera vez,* digamos, dentro de un mes, por influencia de ciertos estímulos ambientales que empezarán a actuar sobre nosotros (por lo que sabemos de nuestro entorno físico o económico) dentro de, digamos, tres semanas.

Ahora bien, hay algo absurdo en esta forma de exponer la cuestión. Porque, si supiéramos hoy, puede argüirse, qué teorías se nos iban a ocurrir por primera vez dentro de un mes, entonces la teoría, naturalmente, se nos ocurriría hoy, en un sentido u otro, y no dentro de un mes; en consecuencia, no habríamos previsto nada que pudiera describirse como el aumento futuro del conocimiento.

Contra este argumento, que yo considero válido, puede formularse la siguiente objeción. Podemos predecir hoy que a alguien se le ocurrirá una idea dentro de un mes, y que sólo entonces será conocida e influyente: y la predicción de hoy debe mantenerse secreta. Pero esta objeción entraña que hayamos predicho el sistema *desde fuera* en vez de desde dentro; porque tomamos medidas (secreto) para no influirlo. Así pues, no era una predicción sobre «nosotros mismos». Además, incluso suponiendo que nosotros mismos pertenezcamos al sistema sobre el que hemos hecho una predicción, sólo podríamos *decidir* mantener los resultados secretos, y no debemos suponer ingenuamente que podamos predecir, científicamente, que llevaremos a cabo nuestras decisiones

—especialmente si cambian las circunstancias—, debido a un aumento inesperado de nuestro conocimiento. La suposición de que podemos predecir tales cosas de nosotros mismos equivaldría a dar por sentada la cuestión que nos ocupa; si es posible, la autopredicción.

Otra objeción parece diferente, pero al final resulta similar. Es que podemos predecir el aumento del conocimiento sin *entender* lo que predecimos. Podemos predecir, digamos, las formas negras que un escritor va a escribir sobre un papel blanco, y sus repercusiones sobre la historia, sin entender todo, o nada siquiera, de lo que intenta expresar. En este caso, no puede decirse que hayamos anticipado las teorías prediciéndolas.

La respuesta es, una vez más, que si podemos predecir, es decir, describir esas formas, nosotros, o cualquiera que esté al tanto de nuestras predicciones, podemos ponerlas por escrito ahora; y, si su génesis va a influir sobre la historia del futuro, no hay ninguna razón para que no lo haga ahora. Debemos admitir que pueden tener diferentes efectos en diferentes circunstancias; pero esto no nos importa ahora. Lo que nos importa en este momento es que no tiene ningún sentido decir que podemos predecir la génesis de nuevas ideas desde dentro del sistema [19].

Y llego ya a la cuestión más importante de la predicción de la *aceptación* de una teoría bajo la influencia de nueva evidencia.

Para no caer en los mismos problemas de antes, deberíamos tener que suponer que la nueva evidencia en cuestión no nos es accesible ahora. En otro caso, nuestra predicción equivaldría a señalar que ahora hay evidencia en favor de una teoría aún no aceptada, y que la teoría debería, por derecho propio, ser aceptada ya. Dicho de otro modo, la predicción volvería a no ser sobre el aumento futuro del conocimiento, sino, más bien, sería un enunciado sobre lo que sabemos ahora.

---

[19] Mi argumento hace uso del hecho de que toda predicción de un sistema desde dentro del mismo puede influirlo; la predicción de lo que, de otro modo, sería una idea nueva cambia su carácter de novedad. He usado el término «efecto-Edipo» como nombre para la influencia, sobre el suceso predicho, de una teoría, o de una expectación, o de una predicción, o incluso de la información utilizada para obtener una predicción: se recordará que fue en parte la influencia de la predicción del oráculo la que hizo que ocurriese el suceso predicho. *Cf.* mis artículos sobre «Indeterminismo», *op. cit.*, pp. 188 ss.; *La miseria del historicismo [Poverty]*, sección 5, y «Philosophy of Science: a Personal Report», en *British Philosophy in the Mid-Century*, 1957, segunda nota a pie de página. [Este último ensayo está reimpreso como «La ciencia: conjeturas y refutaciones», en *El desarrollo del conocimiento científico (C. & R.). Ed.*]

Sería necesario, por tanto, suponer que podemos predecir —sobre la base de nuestro conocimiento actual, es decir, de teorías aceptadas ahora— sucesos aún no observados que, cuando sean observados, proporcionarán evidencia en apoyo de una teoría aún no aceptada, y así llevarán a su aceptación.

Pero eso es imposible. Una evidencia cuya ocurrencia puede predecirse sobre la base de nuestro conocimiento actual no puede ser evidencia que justifique la aceptación de una teoría nueva. Porque la evidencia que pudiera predecirse con la ayuda del conocimiento actual, o no sería nueva o, si lo fuese, equivaldría a una contrastación que confirmase nuestras teorías actuales (en lugar de inducirnos a aceptar una teoría nueva). El tipo de evidencia que justifica la aceptación de una nueva teoría es evidencia que puede predecirse con ayuda de la nueva teoría, pero *no* con ayuda de nuestro conocimiento actual; en otras palabras, tiene que tener el carácter de evidencia *crucial.*

Para mí, este argumento no carece de interés. A pesar de una cierta trivialidad —porque apenas dice nada más que toda teoría entraña su propia verdad [20], y que, por tanto, no puede predecir una situación que implica su rechazo— es suficiente para refutar la influyente doctrina del historicismo; porque muestra que no podemos, por procedimientos científicos, predecir el aumento de nuestro conocimiento teórico. (Podríamos, *en el mejor de los casos,* predecir, en cualquier momento, que nuestro conocimiento no está aumentando ya que nuestras teorías actuales son todas verdaderas y completas.)

Todo esto deja todavía abierta una importante cuestión. ¿Qué pasa si *suponemos* que el aumento de nuestro conocimiento teórico ha llegado a su fin y que nuestras teorías son verdaderas, además de completas? Esto todavía permitiría un cierto tipo de aumento; porque quedaría todavía la tarea ingente de aplicar nuestras teorías a condiciones iniciales siempre nuevas y siempre diferentes. Así, se plantea la pregunta: si fuésemos demonios laplacianos, en el sentido de que conociésemos todas las leyes universales y también todas las condiciones iniciales pertinentes que se aplicarán a nosotros mismos, ¿podríamos entonces predecir nuestras propias predicciones?

---

[20] Estoy refiriéndome al trabajo de Tarski sobre el concepto de verdad según el cual todo enunciado (o, más precisamente, su traducción a un metalenguaje semántico) es equivalente a la afirmación metalingüística de que es verdad.

22. LA IMPOSIBILIDAD DE LA AUTOPREDICCION

Llegaremos así a la última, la más decisiva y la más abstrusa de las cuestiones relacionadas con la predictibilidad del aumento del conocimiento. Es la siguiente:

Suponiendo que estuviéramos pertrechados de un conocimiento teórico perfecto y de unas condiciones iniciales pasadas o presentes, ¿podríamos predecir entonces, por *métodos deductivos,* nuestros propios estados futuros en cualquier instante de tiempo dado y, más especialmente, nuestras propias predicciones futuras?

Naturalmente, no intentaré probar nada más que la imposibilidad de la autopredicción *científica;* es decir, de *deducir* una autopredicción a partir de teorías universales (que se suponen verdaderas), en conjunción con información inicial verdadera sobre el propio estado inicial de uno mismo. Porque los tipos no científicos de autopredicción pueden muy bien tener éxito. Hay, por ejemplo, la autopredicción basada en decisiones de actuar de una determinada manera. Así pues, yo puedo predecir hoy que daré dos conferencias mañana, o puedo predecir hoy que mañana escribiré una carta a mi amigo Fred que empezará con la predicción: «Te sorprenderá saber que...» Las autopredicciones de este tipo no son científicas, no se basan en teorías universales bien contrastadas en conjunción con condiciones iniciales, sino con un proceso de *«tomar decisiones».* Tampoco pueden sustituirse por predicciones científicas basadas en la ley «siempre que decido escribir una carta, al día siguiente llevo a cabo mi decisión», en conjunción con la condición inicial de que «acabo de decidir escribir una carta a Fred mañana»; si no por otra razón, porque la razón de que no soy un sistema cerrado, lo que significa que las condiciones iniciales del sistema que se han descrito son incompletas: puedo recibir un cable esta noche diciéndome que Fred llega mañana, o que han aparecido otras circunstancias nuevas, pertinentes a mi decisión.

Ahora bien, una vez que suponemos que las teorías científicas y las condiciones iniciales están dadas, y también la tarea de predicción, la deducción de la predicción se convierte en un problema de puro cálculo, que en principio puede ser llevado a cabo por *una máquina de predecir o de calcular:* una «calculadora» o una «predictora», como quiera llamársela. Esto hace posible presentar mi prueba en la forma de una prueba de que *ninguna cal-*

culadora o predictora puede predecir deductivamente los resultados de sus propios cálculos o predicciones.

El motivo de formular nuestro problema en términos de máquinas calculadoras tiene varias ventajas de menor importancia. En primer lugar, al hacerlo así hago una concesión a mis adversarios deterministas (sean «materialistas», o «fisicalistas», o «cibernéticistas»); lo que puede persuadirles, a su vez, a considerar mis argumentos con mayor simpatía. En segundo lugar, me permite dar una refutación del determinismo sin suponer la existencia de *mentes.* Me permite, por tanto, dar una refutación que sea más general que una que tuviera que apoyarse en las peculiaridades de predictores humanos. En tercer lugar, todo lo que se ha dicho sobre máquinas será válido, con unos cambios superficiales, también para los predictores humanos. Una cuarta razón es que el método impone una cierta disciplina a aquellos que lo utilicen. No tiene más que una desventaja, a mi modo de ver: que yo puedo ser comprendido con uno de esos que creen que los hombres son máquinas, a pesar de que no creo nada por el estilo [21].

La principal ventaja de exponer nuestro problema en términos de las capacidades de las máquinas de predecir es, sin embargo, éste: podemos imaginar una máquina relativamente simple que represente una especie de modelo simplificado de una teoría determinista *prima facie,* y que sea positivamente *predictible desde fuera.* (Incluso puede ser una máquina capaz sólo de estados dis-

[21] Véase también la última sección de este capítulo, basada en parte en mi artículo sobre «Indeterminismo», citado más arriba, especialmente en las páginas 193-195 de la segunda parte. Puedo mencionar aquí que creo que, en principio, pueden construirse máquinas capaces de realizar *cualquier tarea especificada que puedan realizar los seres humanos.* El énfasis está aquí en las palabras «cualquier... especificada»: porque puede usarse una especificación, en principio, para construir una máquina. Por esta razón, no estoy preparado para aceptar el desafío del mecanicista, *«¡Especifique con precisión* una prueba que sólo pueda pasarla un ser humano y que sea imposible pasarla, en principio, para una máquina!». *Si especificamos con precisión* las pruebas que hay que realizar para descubrir si estamos frente a un hombre o a una máquina, tenemos que hacer frente a la posibilidad de que se construya una máquina para *esas especificaciones* y pase así la prueba especificada. Esto, sin embargo, no significa que, si nos enfrentásemos a una máquina parecida al hombre, tuviésemos la menor dificultad de encontrar infinidad de pruebas que la máquina no fuera capaz de pasar, especialmente si conociéramos su especificación (o incluso si procediésemos simplemente por el método de ensayo y error), aunque pocos seres humanos fallarían dichas pruebas. Véase también mi artículo «El lenguaje y el problema del cuerpo y la mente», *Proceedings of the XIth International Congress of Philosophy,* 1953, vol. VII, pp. 101 y ss. [También reimpreso en *El desarrollo del conocimiento científico (C. & R.);* véase también *El yo y su cerebro (S. & I. B.). Ed.*]

cretos, de forma que todas las cuestiones relacionadas con el problema de la limitada precisión de las condiciones iniciales pueda ser ignorado.) Al mismo tiempo, puede considerarse a la máquina *como una encarnación perfecta, una materialización física perfecta, del demonio laplaciano.*

Para lograr esto supondremos que la predictora es una máquina del tipo siguiente:

Se habrán incorporado a la predictora: *a)* todas las leyes universales verdaderas de la física, y *b)* todos los métodos pertinentes de cálculo de las matemáticas y de la lógica.

La predictora está construida de tal forma que si, y sólo si, está en un estado determinado —su «estado cero»— puede ser estimulada por una *tarea de predicción.* Entonces se aísla de todo otro estímulo y procede a funcionar hasta que ha completado su tarea averiguando una *respuesta;* es decir, una predicción.

Podemos suponer que la *tarea* de predicción consiste en la descripción del estado inicial de un sistema, o el estado en el «tiempo cero», $t_0 = 0$; debe referirse, además, a un instante de tiempo, t_1, por ejemplo, que es aquel instante de tiempo para el que se tiene que predecir el estado del sistema. La predicción constituirá la *respuesta* de la predictora. Desde luego, en lo que estamos interesados fundamentalmente es en esta respuesta: representará la *adición al saber* que va a lograr la predictora, y, por tanto, «el aumento de su conocimiento».

Como un supuesto bastante poco esencial (pero útil en el sentido de que simplifica ciertas cuestiones), podemos añadir que, tras haber dado su respuesta, la predictora volverá a su estado cero.

Para hacer más concretas nuestras consideraciones, podemos imaginar que la tarea de predicción se le suministra a la máquina en forma de una cinta (la *«cinta de la tarea»)* en la que se hacen unas perforaciones que forman el mensaje cifrado en una clave similar a la de Morse. La respuesta se emite en forma de una cinta similar, la *«cinta de respuesta».* Después de completar su tarea, se puede suponer que la máquina consiste en dos partes principales, es decir: *a)* la propia máquina (en el sentido más estricto), que puede estar en su estado cero, y *b)* la cinta de respuesta emitida.

Los dos supuestos siguientes, A1) y A2), sobre la predictora, son esenciales.

A1) Siempre que la tarea suministrada a la máquina sea suficientemente explícita (es decir, lo suficiente para que el demonio

de Laplace llegue a una predicción), la predictora llegará a una respuesta concreta.

Con este supuesto se trata de asegurar que la predictora sea suficientemente eficaz. El siguiente supuesto trata de asegurar que no sea incorpóreo, sino una máquina física.

A2) La predictora *tarda tiempo* en realizar las diversas operaciones. En particular, habrá un lapso de tiempo entre el instante de tiempo en el que la predictora sea estimulada por la tarea de predicción (inserción de la cinta de la tarea) y el instante del tiempo en el que la predictora empiece a escribir (perforar) su respuesta. Aún más, la acción de escribir (perforar) la respuesta lleva también cierto tiempo.

Este supuesto excluye, por ejemplo, a máquinas que estén no sólo dotadas de un conocimiento *teórico* perfecto, sino que sean además omniscientes, o semiomniscientes, en el sentido de que algunas respuestas estén ya incorporadas a ellas y, por tanto, no necesiten ser calculadas. Podemos decir, desde nuestro punto de vista, que las máquinas así dotadas serían *ad hoc*, incluso aunque fueran capaces de responder no sólo una o dos preguntas de esta forma, sino muchísimas.

A partir de estos supuestos, A1) y A2) pueden probarse fácilmente que, en el caso de una tarea de autopredicción, la respuesta sólo puede estar completa *después* del suceso predicho o, como mucho, al mismo tiempo. Esto es suficiente para establecer nuestro punto, que la predictora no puede predecir el aumento *futuro* de su propio conocimiento.

No obstante, si llegamos a fortalecer un poco nuestros supuestos, podemos probar aún más; entonces podemos probar que el predictor fallará totalmente en su tarea. Los dos supuestos adicionales que necesitamos son los siguientes:

A3) De dos respuestas cualesquiera, emitidas por el predictor, la respuesta más larga llevará más tiempo que la más corta.

El cuarto supuesto es el que es decisivo para nuestra prueba de que el predictor fallará, y tendrá que ser discutido con alguna extensión. Es el siguiente:

A4) Todas las respuestas dadas por la máquina describen explícitamente el estado de algún sistema físico en uno y el mismo lenguaje o clave *standard*, o, en otras palabras, excluimos las respuestas especiales que usen un tipo especial de clave o de lenguaje.

Dos consideraciones diferentes nos llevan a este supuesto. Pri-

mero, debemos asegurarnos de que la respuesta dada por la máquina es explícita porque *en ello está toda la finalidad de la máquina.* Porque, en un sentido implícito, la máquina «sabe» la respuesta en el momento en que se le suministran las condiciones iniciales del sistema bajo investigación. Esto es así porque hemos supuesto que la respuesta está entrañada por las leyes incorporadas en la máquina, en conjunción con esas condiciones iniciales; y, aún más —véase A1)—, que la máquina es capaz de responder correctamente a esta información. Así pues, todo lo que le queda por desempeñar a la máquina *es hacer explícita la predicción implícita.* Así, A4) no hace más que expresar con mayor precisión que la máquina es capaz de realizar su tarea.

Como se ha indicado, A2) tiene la función de excluir, por ejemplo, predicciones *ad hoc,* es decir, las máquinas triviales, pseudo-calculadoras. La necesidad de esta exclusión se hace más clara aún cuando tenemos en cuenta la segunda consideración por la que podemos apoyar A4). Es la siguiente:

Si queremos que nuestro predictor opere como una máquina «científica», es decir, una que deduce sus respuestas a partir de leyes y de condiciones iniciales, entonces está claro que debemos excluir ciertas *disposiciones «ad hoc» para la autopredicción.* Por ejemplo, podemos interpretar cualquier sistema físico que cambie periódicamente como un sistema que se autopredice; así, podemos interpretar la noche como una predicción del día siguiente, o de la noche siguiente, etc. Como caso restrictivo, podemos interpretar incluso un sistema invariable como un sistema que se autopredice. Por ejemplo, una «máquina» que consista en una hoja de papel en blanco, convencionalmente, puede interpretarse que contiene un mensaje predictivo en los siguientes términos: «Mientras no se interfiera conmigo, mi estado físico será, en cualquier tiempo futuro *t,* el de una hoja de papel en blanco» [22]. Ejemplos como éste sirven para recordarnos que sólo estamos interesados en máquinas «científicas» o predictoras *deductivas* que no estén designadas *ad hoc* para la autopredicción, sino que sean capaces de predecir deductivamente como mínimo una amplia clase de sistemas físicos diferentes (y entre ellos, si es posible, sistemas muy similares a ellas mismas). No debemos olvidar que estamos interesados por el problema de la autopredicción, principalmente

[22] Debo la idea de esta «máquina» a una comunicación del difunto doctor A. M. Turing recibida alrededor de 1950.

porque fundamentalmente *estamos interesados por el problema de si un predictor puede predecir cambios en aquellas partes de su propio entorno con las cuales él interactúa intensamente.* Pero esto implica que estamos interesados solamente en predictores con una capacidad de predicción muy generalizada, que va mucho más allá de los métodos *ad hoc* de autopredicción del tipo que acabamos de describir.

Ahora bien, el uso de tales *predictores «ad hoc»* estaba excluido implícitamente por nuestro supuesto A2); la que, sin embargo, no estaba excluida por A2) era la adopción de *métodos «ad hoc» de interpretación,* en relación con predictores que satisfagan A2). Por ejemplo, podemos estar de acuerdo en que un predictor bastante «normal» en otros casos —es decir, una que responda a A1) y A2)—, si está en su estado cero, debe interpretarse que expresa el siguiente mensaje: «Soy un sistema físico en tal y tal estado (aquí tendríamos que insertar la descripción física de este predictor en estado cero) y estaré en todo momento en este estado a menos que se me estimule por medio de una tarea.» Esta interpretación determinada puede quizá considerarse excluida por A2), pero puede haber otras similares (por ejemplo, los sistemas periódicos mencionados más arriba) que no sean excluidas.

Para excluir todos los métodos *ad hoc* de esas características tendremos que exigir que, cuando se trate de una tarea de autopredicción, el predictor siga utilizando *esencialmente los mismos métodos* que aplica a todas las demás tareas. En la forma en que acabamos de expresarla, la exigencia es demasiado vaga (como lo indican las palabras «esencialmente» y «métodos»); al mismo tiempo, es también demasiado generalizada; en cualquier caso, parece más fuerte de lo necesario. Porque resulta que todo lo que necesitamos es el supuesto A4) que limita nuestra exigencia a la *respuesta* y al lenguaje en el que se formula [23].

El supuesto A4) excluye la posibilidad de adoptar la convención (por ejemplo) de que un predictor normal en otros casos deba interpretarse que en su enésimo estado se describe a sí mismo y que predice que de este estado pasará al estado $n + 1$ (al que, de hecho, está a punto de cambiar, y el cual, supuestamente, puede siempre calcularse a partir del estado previo).

---

[23] No hay por qué suponer que *todas las tareas* tienen también que ser formuladas explícitamente en un lenguaje. Podemos suponer que la tarea, con las condiciones iniciales, se le da de alguna forma a la máquina; y no necesitamos preguntar cómo. (Véase también la última nota de esta sección [nota 27].)

Todo esto es lo que suponemos sobre la máquina predictora.

Consideremos ahora dos predictores estructuralmente idénticos. Al predictor número 1 se le llama «Dice», porque va a predecir el estado de la número 2; al predictor número 2 se le llama «Dicho», porque va a ser predicho por «Dice». («Dice» apunta hacia «Dicho», por así decirlo.)

Suponemos que las condiciones iniciales suministradas a Dice como parte de su tarea de predicción describen el estado de Dicho a las cero horas ($t_0 = 0$), y que la tarea de Dice es predecir el estado de Dicho a la 1 en punto ($t_1 = 1$). La descripción del estado inicial de Dicho, tal como se le suministra a Dice, tendrá que incluir una descripción de la tarea de predicción (la cinta de la tarea) por medio de la cual Dicho será estimulado a las cero horas. Así que Dice está ahora intentando calcular el estado de Dicho en el instante de tiempo $t_1 = 1$ en punto, o, lo que es lo mismo, el estado de Dicho después de un lapso de tiempo de una hora de longitud.

Según nuestro supuesto 1), Dice tendrá siempre éxito en la tarea de predecir a Dicho.

Ahora suponemos que ocurre que la tarea encomendada a Dice coincide precisamente con la tarea que se encomendará a Dicho a las cero horas; en otras palabras, la tarea de Dice especifica que Dicho será estimulado, a las cero horas, a predecir un tercer predictor. (Esta suposición se hace para que más tarde podamos interpretar la tarea de Dice como una tarea de autopredicción.) Podemos formular éste como nuestro supuesto B).

B) Al recibir el estímulo de su tarea de predicción, Dice estará precisamente en el mismo estado en que esté Dicho al recibir el estímulo de su tarea de predicción, a las cero horas. (Existen razones serias para dudar que podamos suministrar con éxito a Dice una tarea de predicción que informe a Dice de que Dicho está en un determinado estado S, si S va a coincidir con el propio estado de Dice en el momento de recibir dicha información [24]. Pero, como concesión a mis oponentes, supondré aquí que hemos conseguido suministrar a Dicho una tarea de este tipo.)

Supongamos, en primer lugar, que el tiempo de una hora que hemos escogido es tan pequeño que, a la una en punto, Dicho no habrá comenzado todavía a perforar su cinta de respuesta. (En este caso, está claro que no ha tenido lugar ningún aumento del

[24] *Cf.* la última nota de esta sección (nota 27).

conocimiento.) Podemos probar fácilmente el teorema siguiente, T1):

T1) Bajo las condiciones enunciadas, el período de tiempo que le ha tomado a Dice completar su tarea ha sido más largo que una hora.

La prueba es trivial. Puesto que Dice ha completado su tarea, su respuesta ha sido perforada por completo. Pero, después del transcurso de una hora, no puede siquiera haber comenzado a perforar puesto que Dice tiene que pasar por los mismos estados que Dicho, y en los mismos períodos de tiempo; y, de acuerdo con nuestros supuestos, Dicho no habría empezado a perforar su cinta a la una en punto.

A continuación, supongamos que escogemos las dos en punto en lugar de la una como la hora para la que Dice tiene que predecir el estado de Dicho, y que Dicho ha comenzado a perforar su cinta a las dos en punto, aunque sin completarla. Por razones obvias, obtenemos, en este caso, el teorema T2):

T2) Bajo las condiciones enunciadas, el período de tiempo que Dice tardó en completar su tarea fue superior a dos horas.

La prueba es análoga a la anterior.

Ahora supongamos, finalmente, que elegimos las tres en punto como la hora para la que se ha de predecir el estado de Dicho, un espacio de tiempo que es justo lo suficiente para que Dicho complete su tarea de predicción. Obtenemos el teorema T3):

T3) Bajo las condiciones enunciadas, el tiempo que tardó Dice en completar su ejercicio fue exactamente igual a tres horas.

Esto, a su vez, se desprende del hecho de que Dice y Dicho son dos máquinas idénticas; y es suficiente para demostrar que Dice no puede predecir el aumento de su propio conocimiento en el futuro; porque su respuesta completa llegaría demasiado tarde para ser una predicción, puesto que, en el mejor de los casos, sólo llega *al mismo tiempo que el suceso predicho.*

Creo que este resultado es razonable y convincente y que nuestros teoremas establecen todo lo que es necesario para nuestro propósito: la predicción llegará en todos los casos demasiado tarde para que pueda considerársela una predicción del *aumento futuro del conocimiento* de la máquina.

Este resultado se ha deducido sin usar A3) ni A4). Esto significa que es válido incluso aunque introduzcamos *ad hoc* algún simbolismo especial (pero debe ser uno que consuma *tiempo* si se usa) que haga posible la auto-referencia, y que haga posible de

esta manera que una descripción se describa a sí misma [25]. (Está claro que éste es el único caso en el que el auto-cálculo de su respuesta completa puede ser completado por la máquina, aunque demasiado tarde para servir como auto-predicción.)

Pero, si ahora decidimos usar A3) y A4), entonces puede mostrarse, creo yo, que el auto-cálculo se hace totalmente imposible: no sólo llegará tarde, sino que fracasará por completo.

Este puede mostrarse con bastante facilidad si introducimos otro supuesto más, uno muy simple y bastante convincente: de hecho, un teorema auxiliar o lema. Este lema afirma la descripción, o lenguaje *standard* (con ayuda de, digamos, una cinta perforada), del estado físico de una segunda descripción en lenguaje *standard* (de una segunda cinta perforada), no puede en ningún caso ser más corta que la segunda descripción (la segunda cinta perforada). Este lema parece ser verdad en vista del hecho de que tenemos que describir como mínimo cada uno de los símbolos de la segunda descripción (la posición de cada perforación en la cinta), y que cada una de esas descripciones necesitará, como mínimo, un símbolo [26].

Pero, si se acepta este lema, entonces obtenemos el siguiente teorema T4) que contradice T3) y que, por tanto, muestra que el sistema de nuestros supuestos tiene que ser inconsistente.

T4) Bajo las condiciones del teorema T3) el tiempo que Dice tardó en completar su tarea fue superior a tres horas.

La prueba, una vez más, es bastante sencilla, si se acepta el lema. Puesto que Dice tiene que predecir el estado de Dicho a las tres en punto, tiene que describir: *a)* el estado de Dicho aparte de su cinta (lo que es el estado cero de Dicho), y *b)* el estado de la cinta de Dicho. Pero, según el lema, la descripción de Dice de *b)* sólo será por lo menos tan larga como la cinta que va a ser descrita. Por tanto, la descripción que Dice haga de *a)* y *b)* juntos tiene que ser más larga. En vista de A3) esto establece el teorema.

Ahora bien, puesto que T3) y T4) se contradicen mutuamente, el conjunto de nuestros supuestos tiene que ser inconsistente. Esto significa que, si A2), A3), A4) y el lema se cumplen, todos ellos, o A1) o *b),* tienen que ser falsos. Pero esto significa, además, que *el predictor fracasará en la predicción de su propio estado futuro,* bien porque no pueda completar su cálculo, lo que significa que

---

[25] *Cf.* mi artículo sobre «Indeterminismo», etc., pp. 176 y ss.

[26] Este lema lo formulé en mi artículo sobre «Indeterminismo», etc., p. 177.

falló A1), o bien porque no puede proporcionársele la tarea requerida; es decir, una descripción de su propio estado en el momento de serle suministrada esta descripción [27]. Este resultado depende, naturalmente, del lema y de A3) y A4). Pero, incluso sin el lema y sin A3) ni A4), he mostrado que un predictor puede predecir los resultados de sus propias predicciones futuras al menos no antes de que el suceso «predicho» haya tenido lugar de hecho.

Así pues, no podemos predecir el aumento futuro de nuestro conocimiento.

23. LA REFUTACION DEL DETERMINISMO «CIENTIFICO»

Hemos mostrado que es imposible la autopredicción, incluso aunque fuese posible construir un predictor que poseyese los poderes de un demonio laplaciano, y que funcionase por los principios mecánicos más simples; es decir, un predictor que represente un sistema físico cuyo carácter determinista fuese aceptado sin reservas.

Así pues, nuestra prueba no puede, claramente, usarse para refutar el determinismo. Pero puede usarse para refutar el determinismo «científico» y, con él, cualquier pretensión de basar posiciones deterministas en unos resultados cualesquiera de la ciencia, o en el hecho de que la ciencia obtiene buenos resultados.

Porque, si la autopredicción es imposible, entonces está claro que un predictor no puede predecir los efectos de sus propios movimientos en su propio entorno cercano; es decir, la parte de su

[27] En mi artículo anterior sobre «Indeterminismo», citado más arriba, yo mismo (p. 177) formulé una objeción contra un argumento que es similar a la segunda de las dos pruebas que acabo de esbozar. Esta objeción no tiene fuerza contra mi prueba actual, sin embargo, en vista de los argumentos que ahora se proponen en apoyo del supuesto (A4). Por otro lado, puede formularse una objeción análoga contra el argumento de la p. 189 de mi viejo artículo, que expone la dificultad de suministrar a la máquina una cinta de la tarea que describiera a esta misma cinta: podríamos proporcionar al predictor (en contra de mi argumento anterior) de una cinta auto-descriptiva (descrita en mi artículo, en la p. 177), siempre que la máquina estuviera equipada de un «diccionario» o una «máquina de traducir». Creo, sin embargo, que esto sólo cambiaría la retrogresión desde la cinta de la tarea a la propia máquina. Pero aunque creo que podría probarse, a partir de un conjunto de supuestos razonables, que es imposible satisfacer B), es decir, suministrar al predictor una cinta de la tarea que dé su propio estado inicial en el instante mismo de recibir la cinta, he decidido dejar aquí abierta la cuestión de si es B) lo que falla, o A1).

entorno sobre la que él influye de forma apreciable. Y esto significa, además, que la *predicción desde dentro* no puede realizarse con ningún grado de precisión que se haya estipulado, sino sólo en la medida en que la interacción entre el predictor y su entorno pueda despreciarse.

Este resultado viene apoyado por el éxito de la ciencia: aplicamos el método de la predicción científica sólo a sistemas que no sean afectados en absoluto, o lo sean sólo de manera ligerísima por el proceso de la predicción. Por otro lado, el determinismo «científico» requiere que seamos capaces, en principio, de predecir desde dentro todas las cosas de nuestro mundo con el grado de precisión que decidamos; y, puesto que nosotros mismos estamos en nuestro mundo, esta doctrina es refutada por la imposibilidad de obtener predicciones desde dentro arbitrariamente exactas, lo que es una consecuencia de la imposibilidad de la autopredicción.

Este resultado no puede ser debilitado por ningún intento de operar con más de un predictor, siempre que esos predictores estén *dentro del sistema en cuestión:* un predictor diferente de nuestro número 1 puede predecir las dificultades del número 1; sus estados y su influencia sobre el resto del sistema; pero no será capaz de predecir su propia influencia (por ejemplo, sobre el número 1). Además, una «sociedad» de predictores en interacción mutua [28] puede considerarse siempre, formalmente, como *un* predictor complejo; y nuestro resultado es válido para predictores de cualquier grado de complejidad.

Puesto que su refutación proviene, en efecto, del uso exclusivo de la lógica, el determinismo «científico» resulta ser una doctrina que se contradice a sí misma. Así pues, no hay nada que pueda sostener el determinismo «científico»; y ninguna apelación a ninguna ciencia determinista *prima facie,* por muy completa que sea, puede sostener *cualquier otra forma* de determinismo. Las preo-

[28] La influencia del «conocimiento» emitido (la respuesta) de un predictor sobre otro predictor (o cualquier máquina que actúe sobre información) puede hacerse tan fuerte como se quiera: pequeñas diferencias en la cinta de respuesta —por ejemplo, la transferencia de una cinta de respuesta emitida por una máquina a otra que contenga un agujero perforado de más o de menos— pueden dar lugar a una diferencia tan grande como queramos en la respuesta de esta otra máquina. En consecuencia, una «sociedad» de predictores que se influyan mutuamente y que traten de predecirse mutuamente puede ser sumamente inestable, y los predictores individuales pueden perder eficacia si pertenecen a una «sociedad» de este tipo. Cuestiones similares a éstas se tratan en la sección 7 de mi artículo sobre «Indeterminismo», citado más arriba.

cupaciones de Kant eran, pues, innecesarias y ningún filósofo necesita preocuparse de las dificultades que le plantean sus convicciones éticas a causa de un determinismo basado en el éxito de la ciencia (sea empírica o *a priori*).

Nuestra refutación del determinismo «científico» puede dar la impresión de dejar espacio para que el determinismo «científico» sea reemplazado por una tercera versión no afectada por nuestra crítica (la llamo tercera versión en vista de las dos versiones que se discutieron en la sección 12). Esta tercera versión podría expresarse de la manera siguiente: todo sistema físico es predictible en el sentido de que, *al menos después de que el suceso que iba a predecir ha ocurrido,* podemos ver que el suceso estaba determinado por el estado del sistema, en el sentido de que una descripción suficientemente completa del sistema (junto con las leyes naturales) *entraña lógicamente* la predicción. El hecho de que esta predicción no pueda siempre calcularse de antemano no afecta a la lógica de la situación; como se desprende del hecho de que supusimos, en nuestra prueba, un sistema determinista. Y puede decirse, por tanto, que mi prueba no logra su propósito [29].

Mi réplica a esta crítica es que no comprende el verdadero sentido de mi argumento. No quiero refurar el determinismo que, creo yo, es irrefutable; quiero refutar lo que he llamado determinismo «científico». Es perfectamente cierto que al hacerlo no he refutado la tercera versión que aquí se menciona. Pero he refutado a aquellos que se refieren *al éxito real de las predicciones científicas* y que afirman que ese éxito justifica la suposición de que podemos perfeccionar, en principio, nuestras predicciones hasta hacerlas tan precisas como deseemos. Dicho de otro modo, quiero refutar no sólo el determinismo «científico», sino también a aquellos que dicen que el determinismo (más que el determinismo «científico») está justificado por la experiencia científica, y que no es más que una extrapolación legítima. He refutado este importantísimo argumento en favor del determinismo mostrando que no es válido ni siquiera en un mundo determinista. Está claro que este argumento tiene, por su misma concepción, que ser compati-

---

[29] Un argumento que es muy similar al que se da aquí ha sido propuesto por W. B. Gallie en dos artículos muy interesantes: «The Limits of Prediction» (en S. Körner, ed., *Observation and Interpretation*, 1957, pp. 160 y ss.) y su conferencia inaugural *Free Will and Determinism yet Again* (Belfast, 1957). Quiero reconocer mi deuda hacia el profesor Gallie por atraer mi atención, con su crítica, a una laguna en mi argumento.

ble con otras formas de determinismo, tales como la «tercera versión» sugerida. Pero esto no significa que tengamos ninguna razón para creer que esta tercera versión del determinismo sea verdadera o que el mundo tenga una estructura del tipo descrito por esta tercera versión.

Por el contrario, hay muchas razones para creer que el problema de la capacidad de dar razón es suficiente para rechazar esta tercera versión. Hay muchas razones para creer que no podemos recoger suficientes datos para *entrañar* la solución de nuestra tarea de predicción porque no sabemos qué datos serán necesarios para nuestra predicción, incluso bajo el supuesto de que esos datos «existan» en el sentido de que una descripción verdadera, suficientemente detallada, del estado del sistema entrañaría, junto con las leyes naturales, la predicción. Pero a esto podría responderse con una cuarta versión: que una descripción verdadera y suficientemente detallada, si fuera obtenible, entrañaría siempre la solución de cualquier tarea de predicción.

Pero está claro que, por lo menos, esta cuarta versión es completamente metafísica. Opera con un supuesto existencial que es, en principio, irrefutable: el supuesto de la *existencia* de una descripción verdadera que no sabemos cómo obtener.

En resumen, no era mi objetivo refutar aquí la «tercera versión», que no parece refutable por *pura lógica,* puesto que un mundo mecánico muy simple parece posible lógicamente. Mi objetivo era simplemente mostrar que el indudable éxito de la predicción científica no debe usarse como argumento en apoyo de una cosmología determinista: en apoyo de la conjetura de que *nuestro mundo* es del carácter sugerido por la «tercera versión».

La forma en que se ha refutado aquí el determinismo «científico» me parece bastante interesante. No sólo muestra que *no podemos sustituir nuestras decisiones por predicciones científicas* respecto a nuestras acciones futuras (ya que las predicciones de esta clase son imposibles); también muestra que el argumento decisivo en favor del indeterminismo está en la existencia del conocimiento regional mismo. Somos «libres» (o como se quiera llamarlo), no porque estemos sujetos a la suerte en lugar de a unas leyes naturales estrictas, sino porque la racionalización progresiva del mundo —el intento de atrapar el mundo en la red del conocimiento— tiene límites, en cualquier momento, en el aumento del

conocimiento mismo que, naturalmente, es también un proceso que pertenece al mundo.

La acción racional sin algún conocimiento previo —de tipo científico o, como mínimo, hipotético— es imposible; y es ese mismo conocimiento previo el que resulta ser tan limitado como para dejarnos espacio para la acción, es decir, para la libre acción.

24. UN ARGUMENTO DE SAN AGUSTIN, DESCARTES Y HALDANE

Me parece a mí que nuestra refutación del determinismo «científico» guarda cierta relación con un argumento de J. B. S. Haldane. Argumentos similares, pero algo más débiles, fueron propuestos por Descartes y, antes todavía, por San Agustín [30].

El núcleo del argumento es, en la versión de Descartes, que *una comprensión crítica de la verdad,* y una valoración adecuada de un argumento, tiene que ser una acción libre, voluntaria, nuestra (y no la acción de una máquina grabadora); como indica Descartes, la perseverancia en el error y en el prejuicio puede consistir en un fracaso o un rechazo a actuar así libremente —puede consistir en permitirse a sí mismo ser determinado acríticamente por, digamos, las enseñanzas recibidas y permanecer de esta manera bajo el dominio del adoctrinamiento.

Una idea similar fue expresada clara y brillantemente por Haldane; no como crítica del determinismo, es cierto, sino como crítica del materialismo. «Yo no soy materialista —escribió Haldane—, porque, si el materialismo es verdadero, me parece que no podemos saber que lo es. Si mis opiniones son el resultado de los procesos químicos que tienen lugar en mi cerebro, están determinados por las leyes de la química, no por las de la lógica.» Es obvio que lo que critica Haldane aquí no es sólo la idea de materialismo (que es históricamente la versión más importante del determinismo «científico»), sino más bien la idea del propio determinismo «científico». Porque no importa si nos referimos a las

[30] J. B. S. Haldane, *The Inequality of Man,* 1932 (reimpreso Pelican Books, 1937, p. 157); Descartes, *Principles and Philosophy,* parte I, secciones 36-39 (*cf.* especialmente el final de la 37); San Agustín, *De libero arbitrio,* libro I, especialmente los capítulos 11 y 12; libro II, especialmente los capítulos 2, 3 y 19. [Véase la discusión posterior de Popper de esta cuestión en *El yo y su cerebro (S. & I. B.)*, donde también se trata de la retractación de Haldane. *Ed.*]

leyes de la mecánica o de la química, o a las leyes naturales en general. La conclusión es la misma: si mis opiniones están totalmente determinadas por las leyes naturales y por las condiciones iniciales, entonces no están determinadas por las leyes de la lógica. (Aquí, y en el pasaje de Haldane, «lógica» no significa solamente lógica formal, sino algo como el arte de argumentar, y de pesar racionalmente la evidencia.)

En contra del argumento de Haldane puede decirse que una máquina calculadora mecánica, a la que podemos suponer determinada en su funcionamiento por las leyes de la física, puede, no obstante, operar de acuerdo con las leyes de la lógica. Tampoco parece haber ninguna dificultad fundamental para construir máquinas que recojan evidencia empírica y actúen sobre ella (cualquier termómetro que se auto-registre recoge evidencia observacional y cualquier termostato actúa de acuerdo con dicha evidencia). Esto parece refutar el argumento de Haldane.

Pero esta refutación aparente no capta el argumento de Haldane. En cualquier caso, no es aplicable a un argumento similar que yo publiqué no hace mucho [31]. Mi argumento se basa en la distinción entre cuatro funciones diferentes del lenguaje humano: 1) *la función expresiva,* o el lenguaje considerado como síntoma del estado del organismo; 2) *la función de enviar señales,* o el lenguaje como estimulante de respuestas en otros organismos; 3) *la función descriptiva,* o el lenguaje como descriptivo de estados de las cosas (sean existentes o no), y 4) *la función argumentativa,* o el lenguaje considerado como medios de crítica racional (en oposición, por ejemplo, a la simple contra-afirmación [32]). Las dos primeras son también funciones de los lenguajes animales. Las dos últimas pueden ser denominadas funciones «superiores»: ellas dan lugar a la idea de una descripción *verdadera o falsa,* y de un argumento *válido* o *no válido* (creo que podemos distinguir todavía más funcio-

---

[31] *Cf.* «El lenguaje y el problema del cuerpo y la mente», *Proc. of the XIth International Congress of Philosophy,* 7, 1953, pp. 191 y ss. Véase también mi artículo «Nota sobre el problema del cuerpo y la mente: réplica al profesor Wildfred Sellars», en *Analysis,* N. S., 15, 1955, pp. 131 y ss. [Los dos ensayos se han recogido en *El desarrollo del conocimiento científico (C. & R.). Ed.*]

[32] Las tres primeras de esas funciones fueron distinguidas por Karl Bühler, *Sprachtheorie,* 1934, pp. 25-28; yo he añadido la cuarta en «Hacia una teoría racional de la tradición», *The Rationalist Annual,* enero 1949. [Reimpreso en *El desarrollo del conocimiento objetivo (C. & R.).* Véase también la exposición de Popper de esas cuatro funciones del lenguaje en *El yo y su cerebro (S. & I. B.),* y en *Búsqueda sin término (V. Q.). Ed.*]

nes del lenguaje, tales como la prescriptiva o la notificativa o la exhortativa).

Mi tesis, brevemente, es ésta. No podemos, obviamente, describir o argüir sin expresar y enviar señales también. Pero aunque las funciones descriptivas y argumentativa implican las dos funciones inferiores, no obstante, no son reducibles a ellas.

Con esto quiero decir que aunque podamos muy bien decir que describir es una manera especial de expresarse a sí mismo, y de enviar señales, no es *sólo* eso. Porque la *verdad de una descripción* es algo diferente de, digamos, *lo adecuado de una expresión*, o de una reacción a un estímulo; y es también diferente de *lo adecuado de una señal* a una cierta situación, o de su eficacia en provocar la respuesta apropiada a la situación. Porque una descripción puede ser verdadera de hecho, aunque esté hecha con la intención de engañar o de disimular, y puede ser verdadera de hecho, aunque nunca la haya creído nadie e incluso aunque nunca consiguiese provocar la respuesta adecuada.

Tampoco la función argumentativa es reducible a las dos funciones inferiores —expresar y enviar señales—, aunque todos los argumentos expresen y envíen señales. Porque la validez de un argumento no puede reducirse a, digamos, su eficacia persuasiva (que sería una reducción a eficacia en el envío de señales): un argumento válido puede no convencer a nadie; de la misma manera que argumentos no válidos han convencido a mucha gente durante muchos siglos.

Así, la descripción y la argumentación no son simples expresiones y señales. Pero una teoría determinista del lenguaje en términos de las leyes naturales puede explicar solamente esas dos funciones inferiores: tiene que concebir todo lenguaje como sintomático y todas las respuestas a él como respuestas a señales. Lo mismo es válido para cualquier teoría que apele a las máquinas. Una máquina calculadora reacciona a las señales que recibe y los resultados que calcula son expresiones, o síntomas, de sus estados internos. Desde el punto de vista del determinismo «científico», no pueden ser ninguna otra cosa. La diferencia entre una máquina que utiliza un método *válido* de cálculo o de argumento y una que usa un método no válido está más allá del horizonte de cualquier teoría que se confirme a sí misma al enfoque causal del determinismo «científico».

El determinismo «científico» está, por tanto, destinado a ignorar la diferencia entre las funciones «superiores» y las funciones

«inferiores» del lenguaje, o a afirmar la reductibilidad de las funciones superiores a las inferiores; pero esos dos caminos son inaceptables, especialmente porque no hacen justicia a la función y a la estructura del argumento.

Estas concepciones me parecen libres de carácter metafísico o *ad hominem* peculiar de los razonamientos de Descartes y de Haldane; pueden ser defendidas ampliamente (aunque no pretendo hacerlo aquí) basándose en argumentos racionales, como parte de una teoría no-metafísica del lenguaje. Sin embargo, ellas me llevaron a argumentos similares a los de Descartes y Haldane; después de afirmar que las dos funciones superiores no pueden ser explicadas por «filosofías como el conductismo [...], el epifenomenalismo, el paralelismo psicofísico, las soluciones de —dos— lenguajes, el fisicalismo y el materialismo» —es decir, por teorías que mantienen «la completud causal del mundo físico»— añadí: «Todas ésas son *auto-destructivas* en la medida en que establecen —de forma no intencional, por supuesto— la *no-existencia de los argumentos*» [33]. Está claro que puede decirse lo mismo del determinismo «científico». A partir de la capacidad de la razón humana de predecir el mundo —una capacidad que el determinismo «científico» concibe como ilimitada en principio— no deja en su conclusión espacio para el argumento racional, para nuestro discernimiento entre verdad y falsedad, para la diferencia entre lavado de cerebro y aprendizaje.

Este es precisamente el argumento de Haldane. Es la afirmación de que, si el determinismo «científico» es verdadero, no podemos, de manera racional, saber que es verdadero; lo creemos, o no lo creemos, pero no porque juzguemos libremente que los *argumentos o razones* en su favor con correctos, sino porque estamos determinados (o programados) para creerlo, o no creerlo, o incluso para creer que lo juzgamos y lo aceptamos racionalmente.

Este argumento, algo extraño, no refuta naturalmente la doctrina del determinismo «científico». Incluso si se acepta como válido, el mundo puede seguir siendo tal como lo describe el determinismo «científico». Pero, al señalar que si el determinismo «científico» es verdadero no podremos saberlo ni exponerlo racionalmente, Haldane ha dado la refutación de la idea de la que surge el determinismo «científico». No ha refutado la doctrina

[33] «El lenguaje y el problema del cuerpo y la mente», 3.6 (*loc. cit.*, pp. 103 y ss.; las cursivas no están en el original).

pura y simple, pero ciertamente ha refutado *la doctrina en conjunción con su trasfondo filosófico.* La ha refutado en su espíritu, por así decirlo, que es el espíritu del racionalismo o del humanismo científico. Porque Haldane ha mostrado, creo yo, que el determinismo «científico» reduce la racionalidad a una ilusión. El determinismo «científico» entraña la auto-refutación de una visión excesivamente optimista de la razón humana.

Mi argumento desde la imposibilidad de la auto-predicción, expuesto en las secciones precedentes, y el argumento de Haldane tienen en común lo siguiente: los dos operan con la idea de racionalidad y tratan de mostrar que hay una dificultad lógica para considerar que la racionalidad está predeterminada y que es racionalmente predictible. En otros aspectos, los dos argumentos son algo diferentes.

CAPITULO IV

CUESTIONES METAFISICAS

25. LAS DOCTRINAS METAFISICAS DE DETERMINISMO E INDETERMINISMO

Al mostrar la imposibilidad de predecir el aumento del conocimiento, no he mostrado más que la imposibilidad de una predicción completa *desde dentro* del mundo. Esto deja abierta la posibilidad de que el mundo, con todo lo que hay en él, esté completamente determinado *desde fuera,* quizá por la Divinidad. La doctrina metafísica del determinismo deberá ser examinada ahora con mayor atención.

¿Es argüible el determinismo metafísico? Yo creo que sí. A primera vista, uno puede sentirse tentado a decir que el argumento de Haldane (o mi argumento desde las cuatro funciones del lenguaje) establece lo contrario, puesto que muestra que el determinismo no deja espacio para la racionalidad. Pero esto delataría tres errores.

El primer error es éste: el argumento de Haldane no se aplica necesariamente al determinismo religioso (y, por tanto, tampoco a todas las formas de determinismo metafísico); porque de la misma manera que un maestro, ayudado por su superior capacidad de razonamiento, puede ser capaz de predecir el resultado de las deliberaciones racionales de un niño sin reducir por eso la racionalidad del niño a una ilusión, igual podría Dios tener conocimiento previo de nuestras decisiones racionales sin destruir con ello su racionalidad. Es sólo la idea de predeterminación por *las leyes de la naturaleza* (que no tienen capacidad de raciocinio) lo que resulta incompatible con la racionalidad. Así pues, algunas de las formas de determinismo metafísico se ven afectadas por el argumento de Haldane, pero no todas.

El segundo error es éste: el argumento de Haldane muestra que algunas formas de determinismo metafísico entrañan el carácter ilusorio de la razón. Por esto crea dificultades sólo a aquellos que

creen que existe una diferencia entre argumentación razonada y, digamos, indoctrinación sin crítica; es decir, para los racionalistas. Para ellos, la dificultad puede ser insuperable, pero otros deterministas metafísicos pueden ver en el argumento de Haldane una simple refutación del racionalismo en vez del determinismo.

En tercer lugar, no debe interpretarse que el argumento de Haldane establece la irracionalidad de cualquier doctrina determinista, o la imposibilidad de discutir tal doctrina racionalmente; por el contrario, prueba por su propia existencia que es posible argüir sobre el determinismo, porque, ciertamente, es un argumento en contra de él. De la misma manera, mis dos primeras tentativas, en esta sección, de defender el determinismo metafísico contra el alegato de que no es argüible muestran que existen no sólo argumentos en contra del determinismo metafísico, sino también argumentos en su defensa.

Así pues, el determinismo metafísico es ciertamente argüible, pero los argumentos en pro o en contra de él no pueden llegar a ser conclusivos: los que lo apoyan tienen que ser inconclusivos, porque es imposible demostrarse la inexistencia de un suceso indeterminado en el mundo. (Aquí, la situación lógica es similar a la de cualquier teoría universal.) Y los argumentos en su contra tampoco pueden ser concluyentes, porque, por ejemplo, no podemos probar la inexistencia de un espíritu que obtenga pleno conocimiento previo sobre el mundo desde fuera del mundo [1].

Así, tanto el determinismo metafísico como el indeterminismo metafísico son irrefutables. ¿Cómo pueden entonces argüirse sus casos?

En el pasado, el principal argumento en favor del determinis-

---

[1] El determinismo metafísico en su forma religiosa podría formularse por medio de la aserción: «Existe un espíritu que conoce ahora todo acontecimiento del futuro»; y en lo que pudiera llamarse su forma cuasi-científica (es decir, sin la exigencia de predictibilidad desde dentro) por la aserción: «Existe un sistema teórico que, en conjunción con una descripción verdadera del estado presente del mundo, entraña todo conocimiento del futuro». (Hay, naturalmente, otras maneras de formular esta idea.) El indeterminismo metafísico puede expresarse de forma similar: «Existen sucesos en el futuro, tales que ningún espíritu los conoce de antemano y no hay ningún sistema teórico que, en conjunción con una descripción del estado presente del mundo, los entrañe.» Puede observarse que, si expresamos cualquiera de esas teorías en una forma «$(Ex)(y)\ F(x, y)$», su negación tendrá la forma «$(x)(Ey)\ \bar{F}(x, y)$»; pero esta última fórmula puede reforzarse en $(Ey)(x)\ \bar{F}(x, y)$ —como indican mis formulaciones verbales— sin cambiar, naturalmente, su carácter metafísico; además, la formulación más fuerte puede tener ciertas ventajas intuitivas; como de hecho tiene en nuestro caso.

mo metafísico se basaba o bien en fundamentos religiosos o en la fe en el determinismo «científico». Al criticar los argumentos en favor del determinismo «científico» he minado, indirectamente, el determinismo metafísico. Además, el argumento de que el determinismo debe soportar la carga de la prueba (por «prueba» no entiendo una prueba concluyente, naturalmente) es aplicable no sólo al determinismo científico, sino también a la versión metafísica. Y lo mismo ocurre con algunos de mis argumentos filosóficos, por ejemplo el de la asimetría entre el pasado y el futuro; o el argumento de Haldane, aunque no sea aplicable a todas las variantes del determinismo metafísico, tales como el determinismo religioso. Ninguno de estos argumentos son conclusivos. Pero su impacto puede, no obstante, hacerse sentir.

26. POR QUE RECHAZO EL DETERMINISMO METAFISICO. UNA CONVERSACION CON PARMENIDES

En vista de nuestra discusión sobre la relatividad especial, puede uno preguntarse cómo pudo el propio Einstein haber sido un determinista convencido. La respuesta es que, aunque él puede haber creído en el determinismo «científico» en sus años de formación, más tarde, en su vida, su determinismo era de un carácter francamente religioso y metafísico [2].

El vio con claridad que no había ningún argumento válido que condujese del experimento a la teoría; y, sin duda, vio, con la misma claridad, que no había argumento válido que condujese de la ciencia a la metafísica. Pero sus argumentos fueron en la dirección opuesta. El no basó su determinismo metafísico en el carácter determinista *prima facie* de sus teorías físicas, sino que exigió que sus teorías físicas tuviesen ese carácter [3], porque él creía que la propia realidad física era determinista. (De la misma manera, exigía que nuestras teorías fuesen simples, porque creía en la simplicidad del mundo, de la realidad física.)

Einstein encontró interesantes mis argumentos contra el deter-

---

[2] [Pero véase la nota 2 del capítulo I de este tomo del *Post Scriptum*, donde se observa que Einstein aparentemente abandonó el determinismo antes de morir. *Ed.*]

[3] [Véase la nota 2 del capítulo I de este volumen, donde Pauli informa de que en 1954 Einstein ya no exigía que sus teorías fuesen «rigurosamente deterministas». *Ed.*]

minismo «científico» y sintió que abordaban el problema desde un ángulo que él no había tenido en cuenta anteriormente. Pero sintió que, aunque mis argumentos contra el determinismo «científico» fuesen válidos, no debilitarían su determinismo metafísico ni su preferencia por teorías deterministas *prima facie*. Por tanto, intenté un ataque más directo a su determinismo metafísico.

Lo hice en una conversación privada al día siguiente de leer un artículo sobre el mismo tema [4]. Primero traté de escribir su propio determinismo metafísico, y él estuvo de acuerdo con mi descripción. Le llamé Parménides, ya que él creía en un universo cerrado de cuatro dimensiones, inmutable, como el universo cerrado de tres dimensiones de Parménides [5]. (La cuarta dimensión era el tiempo, naturalmente.) El estuvo totalmente de acuerdo con esta descripción de sus ideas y con la analogía de la película: a los ojos de Dios, la película estaba allí de la misma manera que el pasado: nunca pasaba nada en este mundo y el cambio era una ilusión humana, al igual que la diferencia entre el pasado y el futuro [6].

Ataqué esta concepción con dos argumentos:

El primero era que nada en nuestra experiencia de este mundo justificaba una metafísica parmenídea de este tipo. Einstein lo admitió, aunque seguía sin impresionarle demasiado hasta que le recordé que él mismo había utilizado, bastante recientemente, un argumento análogo —que nada en nuestra experiencia justificaba la introducción de acción a distancia— contra un intento de salvar cierta interpretación de la teoría cuántica [7].

Mi segundo argumento era de carácter más metafísico; era éste: si se supiera que el universo estuviera predeterminado, como

---

[4] Tuvo la amabilidad de asistir a la lectura de este artículo en Princeton en 1950 (en la línea de mi artículo sobre «Indeterminismo» al que se refiere la primera nota de este capítulo). [Véase el relato que hace Popper de este encuentro en *Búsqueda sin término (U. Q.)*, pp. 172-177. *Ed.*]

[5] *Cf.* mi artículo «La naturaleza de los problemas filosóficos y sus raíces en la ciencia», *The British Journal for the Philosophy of Science*, 3, número 10, el último párrafo en las pp. 141 y ss. [Reimpreso en *El desarrollo del conocimiento objetivo (C. & R.), Ed.*]

[6] Naturalmente, la presentación cuatridimensional del mundo no es más que una forma de hablar o un lenguaje equivalente, aunque más conveniente para determinados propósitos que el lenguaje ordinario de tres dimensiones (más tiempo); no tiene *implicaciones* metafísicas necesarias (ni siquiera con respecto al pasado y al futuro). Y, sin embargo, *sugiere* muy particularmente una metafísica determinista, de inmutabilidad a los ojos de Dios.

[7] *Cf.* Albert Einstein, «Quanten-Mechanik und Wirklichkeit», *Dialectica*, 2, 1948, números 7 y 8, pp. 320 y ss.; véase especialmente el último párrafo, p. 323.

la película, y tuviera cuatro dimensiones, como la película (porque, si consideramos que sus escenas representan un mundo tridimensional, podemos tomar *el orden de la secuencia de las escenas* como la cuarta dimensión), entonces se seguirían unas series de consecuencias difíciles de aceptar. Señalé tres de ellas: la primera era que el futuro, al estar entrañado causalmente por el pasado, podía considerarse contenido en el pasado, igual que el pollo está contenido en el huevo, con todos sus mínimos detalles. El futuro se convertía, por tanto, *en redundante. Era superfluo.* No tenía ningún sentido ver una película cuyas escenas, en su totalidad, estuvieran entrañadas de forma estrictamente lógica (en conjunción con una teoría conocida) por la primera escena. Además, esa colosal redundancia era difícilmente conciliable con la idea de Einstein de la simplicidad en su sentido metafísico.

Otra consecuencia era que estábamos obligados a interpretar nuestra forma humana de *experimentar el cambio,* y el flujo del tiempo. Esto tendría que hacerse, una vez más, usando la analogía de la película: experimentamos escenas sucesivas o «láminas de tiempo» (por usar un término que se debe a J. H. Woodger) del mundo a nuestro alrededor, y además su orden sucesivo. Pero esto equivale a decir que la flecha del tiempo es subjetiva, y que el tiempo, tal como lo experimentamos, es una ilusión: una idea que forma parte integral de una filosofía idealista o subjetivista y que está vinculada con otras consecuencias idealistas y subjetivistas más. Y una de las convicciones más profundas de Einstein era su realismo.

La última consecuencia, señalé, tenía toda la apariencia de una contradicción pura y simple. Si estuviésemos experimentando las escenas sucesivas de un mundo inmutable, entonces una cosa, al menos, estaría cambiando realmente en el mundo: nuestra experiencia consciente. Una película, aunque exista ya y esté predeterminada, tiene que *pasar, moverse* a través del proyector (es decir, en relación con nosotros mismos) para producir la experiencia, o la ilusión, de cambio temporal. De la misma manera, nosotros tendremos que movernos en relación con el universo cerrado de cuatro dimensiones, porque la conversión de nuestro futuro en nuestro pasado significa un cambio *para nosotros.* Y, como formamos parte del mundo, habrá entonces cambio en el mundo, lo que contradice la idea de Parménides.

Estas críticas, admití, quizá no fueron incontestables, pero no sería fácil dar una respuesta. No serviría de nada mirar a nuestra

propia conciencia extendida en el tiempo y coexistiendo en el tiempo: tendríamos, una vez más, que explicar por qué no se experimentaba de este modo y sí, en cambio, como una sucesión temporal de «láminas de tiempo». El cambio era real y no podía explicar lo contrario sin adoptar una concepción realista del mundo; sin distinguir con Parménides entre una realidad, que no cambia, un mundo ilusorio de apariencia, que sí cambia. Y aun entonces tendríamos que explicar el hecho objetivo —la realidad— de la ilusión, y nuestra incapacidad de deshacernos de ella incluso aunque aceptemos su carácter ilusorio. (En caso de la mayoría de las ilusiones ópticas, el saber que estamos sufriendo una ilusión tampoco consigue disiparla: la ilusión es un hecho, y, efectivamente, un hecho que en muchos casos puede explicarse fisiológicamente.)

En vista de todas esas dificultades, señalé que el camino más simple de todos era rechazar cualquier concepción metafísica que no permitiese la asimetría entre el pasado y el futuro y aceptar una concepción que permitiese que el futuro no estuviera entrañado y, en cierto modo, contenido en el pasado; dicho de otro modo, aceptar una concepción indeterminista del mundo. La metafísica indeterminista parecía estar más cercana a la experiencia y no parecía crear nuevas dificultades de ningún tipo, una vez que se hubiera mostrado que los argumentos en favor del determinismo «científico» no eran válidos.

Esos fueron mis argumentos. Parménides los discutió con gran paciencia, como era su costumbre. Dijo que le habían impresionado y que no encontraba respuesta para ellos. Más allá de esto, no proseguí.

27. EL BENEFICIO PARA LA CIENCIA: UNA TEORIA DE LAS PROPENSIONES

Hasta ahora he criticado el determinismo tratando de mostrar sus desventajas. También he insinuado que el abandonarlo daría lugar a beneficios positivos para el sentido común, la ética, la filosofía de la ciencia, la cosmología y, espero, para la verdad.

Pero en este libro, y en el contexto actual, con la teoría cuántica como uno de nuestros mayores intereses (véase *Teoría cuántica y el cisma en física*, tomo III de este *Post Scriptum*), un beneficio —y quizá uno de la mayor importancia— para la propia cien-

cia lo proporciona quizá el argumento positivo de mayor fuerza en favor del indeteminismo: al rechazar el determinismo abrimos camino a un enfoque que podría ser de auténtica significación para la ciencia. Estoy pensando en una interpretación física de la teoría de las probabilidades bajo la forma de una teoría física de las propensiones. (Véase *Realismo y el objetivo de la Ciencia,* tomo I de este *Post Scriptum,* parte 2.)

Incluso aunque una teoría así, después de una discución seria, resultase finalmente inaceptable, el hecho seguirá siendo que solamente después de descartar el determinismo ganamos la libertad necesaria para una consideración seria de la interpretación propensivista como teoría física. Así pues, el determinismo no sólo carece del apoyo de los argumentos; nos impide considerar seriamente posibilidades —tales como la idea de propensiones físicas— que, sea cual fuere la evaluación final de sus méritos, son ciertamente merecedoras de ser tenidas en cuenta.

La mejor manera de explicar la idea de las propensiones físicas es por medio de una analogía con la idea de las fuerzas físicas. Una fuerza física es una entidad hipotética imposible de observar, pero contrastable; contrastable, naturalmente, mediante la contrastación de una hipótesis que incluya una fuerza. Por ejemplo, la hipótesis de la presencia, en un lugar determinado, de una fuerza electrostática de una cierta dirección e intensidad puede contrastarse por sus efectos predictibles: la dirección y la magnitud de la aceleración de un cuerpo de contrastación que puede colocarse en dicho lugar.

Supongamos ahora que en una secuencia de contrastaciones todas ellas producen los mismos resultados: podemos explicar esto con la hipótesis de que la fuerza es constante. Supongamos, por otro lado, que en una secuencia de contrastaciones todas producen los mismos resultados respecto a la dirección de la fuerza, pero que los resultados referentes a la magnitud de la aceleración fluctúan: entonces lo explicamos con la hipótesis de que la dirección de la fuerza es constante, mientras que su intensidad fluctúa. Una interpretación equivalente puede darse en el caso de que la dirección de la aceleración fluctúe, pero no su magnitud.

Pero en ciertos casos puede ocurrir que esas hipótesis de fuerzas fluctuantes sean teóricamente insatisfactorias; por ejemplo, porque, durante la contrastación, hemos mantenido todas las condiciones tan constantes como hemos podido. En este caso, quizá podamos explicar las fluctuaciones diciendo que son debidas a

perturbaciones desconocidas, o a fuentes desconocidas de fracaso que impiden mantener constantes las condiciones experimentales. Pero también esto puede ser insatisfactorio, y podemos entonces decidir introducir una nueva idea. Podemos decir que la *situación objetiva,* todas cuyas condiciones hemos mantenido constantes, determina *propensiones en lugar de fuerzas;* que determina propensiones a acelerar —o propensiones a resultar acelerado— que pueden ser más elevadas cerca de la media de las aceleraciones e ir variando gradualmente tanto hacia valores más elevados como hacia valores más bajos. La hipótesis de la presencia de tales propensiones tendría que ser contrastada por medio de contrastaciones estadísticas (como se indica en el tomo I, parte 2 del *Post Scriptum,* sobre probabilidad).

Sólo puede introducirse esta idea, naturalmente, una vez que hayamos renunciado al determinismo; porque el supuesto es que *la misma situación puede producir resultados fluctuantes.* Queda claro, después de nuestra discusión, que, si no renunciamos al determinismo, podemos tener que operar con la idea de fuerzas fluctuantes en vez de con la idea de propensiones; y queda claro también que, en ciertos casos, las dos explicaciones pueden resultar matemáticamente equivalentes. ¿Cuál de ellas, entonces, es la que aceptaremos?

La respuesta a esta cuestión no puede darse con ninguna seguridad, pero las cuestiones de contrastabilidad pueden ser decisivas. La interpretación determinista en términos de fuerzas fluctuantes tendrá que postular unas *condiciones iniciales fluctuantes.* Si puede contrastarse este postulado, y resiste las contrastaciones, la explicación del determinista, en términos de fuerzas fluctuantes, vence. Si, por el contrario, el determinista se ve forzado a recurrir a una hipótesis incontrastable sobre *fluctuaciones ocultas de las condiciones iniciales,* una explicación en términos de propensiones, cuya presencia pueda contrastarse estadísticamente, puede resultar preferible. (Otras circunstancias que pueden hacerla preferible serán discutidas en la próxima sección.) En cualquier caso, no debe permitirse que un prejuicio en favor de la teoría del determinismo impida la libre discusión de la teoría de las propensiones.

He intentado explicar la idea de propensión como una especie de generalización —o quizá como una alternativa incluso— de la idea de fuerza, principalmente porque también la idea de fuerza fue vista con desconfianza al principio por los físicos racionalistas

que la calificaron, con toda razón, de oculta y metafísica. Pero desde entonces hemos aprendido (o al menos eso espero) que la ciencia física explica lo conocido por medio de lo desconocido [8] y el mundo visible por medio de un hipotético mundo invisible; y nos hemos acostumbrado ya a la idea de fuerzas. (Newton no estuvo nunca muy satisfecho con la idea de una fuerza de atracción; Heinrich Hertz intentó prescindir de ella, y lo mismo hizo Einstein.) Así pues, nosotros podemos acostumbrarnos, de la misma manera, a la idea de propensiones.

Al hacer una analogía entre propensiones y fuerzas, no intento sugerir que tengamos que considerar solamente, o principalmente, las propensiones a acelerar o a ser acelerado. Por el contrario, puede haber otras propensiones de mayor importancia: en general, consideramos que las propensiones asumen, bajo determinadas condiciones, uno u otro de una serie de estados «posibles» (o «virtuales»).

Los valores numéricos de las propensiones a tomar diferentes estados pueden variar de un estado a otro. La función (la distribución de probabilidad) que determina estos valores reflejará, en general, las simetrías o asimetrías de las condiciones. Puede introducirse una analogía con fuerzas en dos puntos: podemos tener que considerar las propensiones (o quizá otras funciones estrechamente relacionadas con ellas) como magnitudes físicas hipotéticas, capaces incluso de interacción (o quizá interferencia) como las fuerzas. Y podemos tener que atribuir pesos a las propensiones: pesos que, de acuerdo con las simetrías inherentes a la situación, pueden no estar totalmente determinados sólo por esas simetrías.

28. TEORIAS DETERMINISTAS *PRIMA FACIE* Y TEORIAS PROBABILISTAS

Las teorías deterministas *prima facie* no pueden responder a todas las cuestiones que pueden plantearse, legítimamente, en físi-

[8] Para la idea de explicar lo conocido por medio de lo desconocido, véanse mis artículos «La naturaleza de los problemas filosóficos y sus raíces en la ciencia», *The British Journal for the Philosophy of Science*, 3, 1952, pp. 124 y ss., especialmente 144, 148 y ss., y «Tres concepciones sobre el conocimiento humano», en *Contemporary British Philosophy*, III, ed. por H. D. Lewis, 1956. [Los dos están reimpresos en *El desarrollo del conocimiento científico (C. & R.)*. Véase también el tomo I de este *Post Scriptum*, sección 15. *Ed.*]

ca. No pueden responder a preguntas tan simples como ¿por qué una batidora consigue mezclar siempre, digamos, granos de café y granos de cacao que se han metido en ella en dos montones separados?, o la pregunta, muy similar, ¿cómo es que una máquina de tirar a cara o cruz, construida de forma adecuada, en la que siempre metemos los peniques precisamente de la misma manera, produce secuencias de tiradas de peniques de un carácter aleatorio? [9]. Está claro que ésos son problemas físicos que no pueden desecharse; y, puesto que lo que hacen son *preguntas esencialmente estadísticas,* éstas deben ser contestadas por teorías que sean esencialmente estadísticas o probabilistas.

Quizá los más característicos y los más importantes de esos problemas en física sean los de la intensidad de las líneas espectrales y los de la vida media de los núcleos que se desintegran radiactivamente.

En el núcleo de *la lógica de la investigación científica* afirmé en muchas ocasiones que las conclusiones estadísticas (sintéticas) sólo pueden obtenerse a partir de premisas estadísticas. En vista de mi trabajo posterior sobre probabilidades (véase la parte 2 del tomo I del *Post Scriptum),* tengo que reformular eso.

Uso aquí la palabra *«probabilista»* en un sentido amplio, de forma que englobe teorías «objetivas» que o son teorías estadísticas (teorías sobre secuencias) o teorías que afirman una probabilidad en el sentido de probabilidad de la teoría de conjuntos o de la interpretación propensivista. (No tomo en cuenta aquí teorías subjetivas o probabilidades «inductivas» en el sentido de, digamos, Jeffreys, Keynes o Carnap [10].)

En este uso, mi antigua afirmación —de que las conclusiones estadísticas sólo pueden obtenerse a partir de premisas estadísticas— puede ser sustituida por las siguientes:

1) Las conclusiones probabilistas (más pecisamente, los enunciados no analíticos, que afirman probabilidades o frecuencias distintas de 0 ó 1) pueden deducirse sólo a partir de premisas probabilistas.

2) Las conclusiones estadísticas pueden deducirse de premisas

---

[9] Un paso importante en esta dirección fue el que dio Khinchine; véase *Sowjet-wissenschaft,* 1954, pp. 268 y ss. El llama a su método «el método de la distribución arbitraria de funciones». (El Dr. I. Lakatos me llamó la atención sobre este artículo.)

[10] Véase *Realismo y el objetivo de la ciencia,* tomo I de este *Post Scriptum,* parte 2.

estadísticas o de otras premisas probabilistas. Si se derivan de premisas probabilistas no-estadísticas, estrictamente hablando, no se deducen conclusiones estadísticas; pero si interpretamos las probabilidades o medias 0 y 1 como «casi nunca» o «casi siempre», podemos decir que las conclusiones estadísticas «casi se deducen» de premisas probabilistas no-estadísticas.

En vista de 2), algunas veces podemos ser capaces de *contrastar* teorías probabilistas por medio de la contrastación de enunciados estadísticos que *«casi se deducen»* de las premisas probabilistas. Por ejemplo, podemos contrastar, por medio de una secuencia de ensayos, la teoría de que la probabilidad de sacar cara con cierta máquina es igual a ½; porque en el supuesto de que la probabilidad de sacar caras sea de ½, y de que las tiradas sean independientes, de forma que la probabilidad sea constante, la probabilidad de que una secuencia de ensayos no sea una secuencia de tipo aleatoriforme con la frecuencia ½ será cero; de donde se deduce que la mayor parte de las secuencias observadas durante largo tiempo estarán muy próximas a «realizar» las propiedades frecuenciales de una secuencia o colectivo aleatorios.

Este ejemplo sólo pretende recordar al lector cómo los supuestos estadísticos pueden «casi» deducirse de premisas no-estadísticas probabilistas. No pretende ser, de ningún modo, una respuesta a la pregunta, hecha al principio de esta sección, de cómo explicar los efectos observables obtenidos por una batidora o (lo que es equivalente) por una máquina de tirar a cara o cruz.

Puede hacerse una tentativa determinista de responder a esas preguntas, combinando una teoría determinista *prima facie* de los procesos físicos implicados con un supuesto referente a las condiciones iniciales; un supuesto que podía describirse como *un supuesto probabilista sobre condiciones iniciales «ocultas»*.

Para explicarlo más completamente con ayuda de un ejemplo, supongamos que nuestra máquina de tirar a cara o cruz está construida con gran precisión, para que repita, o reproduzca, sus propios movimientos con mucha exactitud. Supongamos que la máquina coge varias veces el penique, lo sostiene en una posición estrictamente vertical, lo hace girar sobre un eje vertical y luego lo deja caer, todavía girando, sobre un plano inclinado sobre el que o bien rueda hacia abajo o se desliza (si ha caído de lado antes de llegar al final del plano). El proceso se repite, digamos, veinte veces, después de las cuales el penique sale.

¿Cómo explicamos el hecho de que una máquina de este tipo

consiga producir una secuencia de peniques muy bien «mezclada» o «aleatoria» en la que aproximadamente en la mitad de los peniques sale cara y en la otra mitad cruz?

Difícilmente podremos atribuir este hecho a una irregularidad en la manera de introducir el penique en la máquina, porque: *a)* si lo introducimos en ella siempre de la misma manera, con la mayor exactitud posible, el resultado estadístico no es afectado, y *b)* si variamos nuestro método de introducción, el resultado estadístico sigue sin ser afectado. Más aún, podemos construir la máquina de tal manera que en los primeros pasos corrija con gran precisión las diferencias de posición del penique que puedan existir al introducirlo en la máquina: la máquina puede así *igualar* las condiciones iniciales (aunque, claro, no totalmente).

En vista de todo esto, podemos sentirnos inclinados a atribuir el resultado estadístico a diminutas y *ocultas diferencias en el estado de la máquina y del penique* —por ejemplo, a cambios atómicos o moleculares—; es decir, podemos atribuir las diferencias en los resultados a diferencias en las condiciones iniciales ocultas. Podemos entonces explicar el diferente resultado macroscópico señalando que la máquina tiene un mecanismo (tal como dejar caer verticalmente el penique, mientras gira, varias veces, en nuestro ejemplo; o en otra máquina, sacudiéndolo violentamente) que equivale a una amplificación de las diminutas diferencias ocultas que inevitablemente ocurren en los diversos movimientos de la máquina.

Esto explica —creo yo que bastante satisfactoriamente— el hecho de que la máquina no produzca siempre caras, y que algunas veces salgan cruces. Pero es insuficiente para explicar la *estabilidad* estadística del resultado el hecho de que la salida observada de la máquina concuerde estrechamente con la hipótesis de que produce un colectivo, con frecuencia relativa ½, digamos [11].

Para explicar esto tenemos que suponer (i) que la secuencia de las condiciones iniciales ocultas forma también un colectivo. Y esto, a su vez, puede explicarse aún más suponiendo (ii) que cualquier supuesto que no sea (i) es sumamente improbable, que el conjunto de las condiciones iniciales que *no* forman un colectivo

[11] La máquina podría mostrar una «predilección» y producir frecuencias que fueran mayores de ½; o podría cambiar periódicamente sus frecuencias entre, digamos, 0,45 y 0,55 (en este caso, los períodos tendrían que ser bastante largos para que pudieran averiguarse).

fortuito tiene la probabilidad o medida 0 [12]. De esta forma, nuestro problema estadístico se resuelve, en último término, por medio de una deducción a partir de un supuesto probabilista, pero no-estadístico, sobre las condiciones iniciales ocultas. O, en otras palabras, nuestro problema estadístico se resolvió por medio de una teoría probabilista; porque la teoría determinista *prima facie* de la máquina tiene sólo un papel muy subordinado en la explicación del efecto *estadístico.*

Considero que una explicación del tipo de la que se ha dado aquí es satisfactoria hasta cierto punto; pero es importante percatarse de que no opera con una teoría puramente estadística, sino con una teoría de la probabilidad o de la medida. Porque hemos explicado el supuesto —designado más arriba como «(i)»— de que las condiciones iniciales constituyen un colectivo fortuito por medio de otro supuesto —designado como «(ii)»— de que la ocurrencia de cualquier otra secuencia tendría una medida o una probabilidad 0. Esto significa, sin embargo, que suponemos que es válida una teoría de medida no estadística de la probabilidad para la distribución de nuestras condiciones iniciales y que *esta* teoría de la probabilidad tiene que interpretarse físicamente (yo sugiero que por propensiones).

Una teoría puramente estadística no serviría. Significaría pararse en (i) y explicar, por tanto, el carácter aleatoriforme de las secuencias de las tiradas (observables) de peniques *suponiéndole* el mismo carácter a la secuencia de las condiciones iniciales ocultas. Pero esto no hace más que hacer retroceder un paso en la cuestión. Además de ello, algunas de esas últimas secuencias pueden muy bien ser de hecho aleatoriformes; pero ¿qué derecho tenemos para predecir (como lo hacemos) que todas o casi todas ellas lo serán por *norma,* por cuestión de *ley*?

29. LA CUCHILLA DE LANDE

No conozco ningún físico que haya visto este problema con más claridad, o que haya hecho más por mostrar lo que hay im-

[12] Así, tendríamos que mostrar, por ejemplo, que casi todos (es decir, *todos* menos un conjunto de medida cero) los estados iniciales de un gas llevan a estados de equilibrio (o a una distribución maxwelliana de las velocidades de las moléculas). Todas ésas son tareas probabilistas y tareas que han sido resueltas o están siéndolo. Pero la solución sólo tiene sentido si interpretamos la medida cero como una propensión cero.

plicado en él, que Alfred Landé. Su argumento [13] está concebido para mostrar que tenemos que aceptar como fundamentales las probabilidades de sucesos únicos, y aceptar que son insustituibles por cualquier enunciado que no sea otro enunciado de probabilidades. Además, su argumento muestra que, incluso aunque combinemos una teoría determinista *prima facie* con supuestos estadísticos referentes a las condiciones iniciales, no obtendremos más que una retrogresión infinita; y una interpretación que persevere en ese supuesto está destinada a convertirse en incontrastable, metafísica (o «puramente académica» en la terminología de Landé). Voy a citar todo el importante pasaje de Landé (que, de paso, contiene también un argumento contra el determinismo):

> Por un tubo se dejan caer unas bolas de marfil en el centro de una cuchilla de acero y se observa que una proporción media de 50 : 50 de las bolas cae hacia la derecha *[d]* y otro 50 : 50 cae hacia la izquierda *[i]*. Ahora bien, aunque un observador superficial puede considerar un suceso *d* individual como puramente accidental, un físico más experto puede ser capaz de ver de antemano que una bola *d*, incluso antes de chocar con la cuchilla, tenía una ligera preponderancia hacia la derecha. Este ver de antemano presupone que el observador tiene un artefacto óptico, una especie de cuchilla óptica que cumple la misma función de distinguir entre bolas *d* y bolas *i*, que la que realiza más tarde la cuchilla de acero. Uno de los sucesos en la vida de una bola *d* puede haber sido un encuentro predestinado con la concepción clásica, entonces, el estado *d* de hoy está precedido por el estado *d* de ayer y así hasta el pasado infinito a través de una cadena ininterrumpida de sucesos ... *ddd* ... de los cuales el encuentro con la cuchilla de acero no es más que un eslabón.
>
> Cuando se le pida ahora al determinista una explicación causal de la proporción media de 50 : 50 entre bolas *d* y bolas *i*, su respuesta será que *también esta proporción estaba predeterminada* mucho antes de que existiesen el tubo y la cuchilla. Si se le presiona más todavía [a explicar] por qué incluso las fluctuaciones de la media concuerdan con las expectativas estadísticas de la teoría de los sucesos aleatorios, puede retirarse a conceder que existe una *armonía pre-establecida* entre grupos de sucesos que *parecen* sujetos a fluctuaciones aleatorias, aunque *en realidad* cada suceso individual estaba predeterminado. Sin embargo, esto pondría al «parecer» y a la «realidad» en una posición invertida. La distribución aleatoria es una realidad física y ... [un sistema determinista] que simplemente parece aleatorio es una construcción puramente académica. Una distribución de los efectos que satisfaga la teoría del error requiere ... desde el punto de vista del

---

[13] Véase Alfred Landé, «Probability in Classical and Quantum Theory», *Scientific Papers Presented to Max Born*, 1953, pp. 58 y ss., y Alfred Landé, *Foundations of Quantum Theory*, 1955, pp. 3 y ss. [Véase también Landé, *From Dualism to Unity in Quantum Physics*, 1960, pp. 3-8, y *New Foundations of Quantum Mechanics*, 1965, pp. 29-32. *Ed.*]

determinista, una distribución de las causas correspondientemente aleatoria en un tiempo anterior y de ahí a un tiempo anterior a ése. El programa de dar una teoría estrictamente determinista de los sucesos distribuidos estadísticamente no lleva a ninguna parte.

El sencillo, pero precioso, argumento de Landé puede expresarse de forma aún más explícita.

a) Supongamos que el número *N* de las bolas que se dejaron caer era 1.000. El determinista, como señala Landé, puede explicar tanto la proporción de 50 : 50 de los resultados como la fluctuación aleatoria, sólo mediante la suposición de que había una distribución correspondiente en las condiciones iniciales de cada uno de los 1.000 sucesos individuales. Si trata de explicar por qué esas 1.000 condiciones iniciales muestran la proporción de 50 : 50 y por qué muestran las fluctuaciones aleatorias, está, claramente, camino de una retrogresión infinita. Si se niega a intentar dar una explicación, tiene que aceptar el carácter inexplicable, milagroso de los hechos.

Pero puede empujársele más lejos: sin duda conjeturará —y si no lo hace él, otros lo harán— que los próximos 1.000 sucesos, o los próximos 10.000 sucesos, producirán resultados estadísticos muy similares. Entonces tendrá que conjeturar que también ellos se deben a una distribución correspondiente de las condiciones iniciales; y será incapaz de decir por qué conjetura el que esas proporciones van a ser tan extrañamente estables. (En este sentido, tendrá que creer de nuevo en una «armonía pre-establecida», como dice Landé.)

Lo que Landé muestra aquí es la futilidad de la vieja «explicación determinista, según la cual muchas pequeñas causas o «errores» producirán (en parte, el neutralizarse mutuamente, etc.) el resultado aleatorio. Todo esto puede ser verdad, ciertamente; pero eso no cambia el hecho de que, para el determinista, *los resultados estadísticos* sean derivables, si es que lo son, sólo de *suposiciones estadísticas* referentes a la distribución de las condiciones iniciales. Así pues, encontramos que la extraña conducta, como regida por una ley, de las secuencias estadísticas permanece, para el determinista, *irreductible e inexplicable, en último término.* Más especialmente, el determinista no puede explicarla como debida a un elemento de aleatoriedad o casualidad, o apelando a elevadas probabilidades: el argumento de Landé muestra que esas ideas se hacen inaplicables, puesto que el determinismo no puede apelar más que a la distribución estadística, inexplicada, de la secuencia

de sucesos anteriores (es decir, de la secuencia de las condiciones iniciales).

Estas consideraciones de Landé sugieren con fuerza que los deterministas que creen que pueden explicar la conducta estadística permiten que consideraciones de probabilidades (e incluso de propensión) se deslicen sin ser notadas, en sus supuestos. Los deterministas operan con un supuesto al que puede denominarse *la hipótesis general de la aleatoriedad:* el supuesto de que las condiciones iniciales que están fuera de control son siempre aleatorias. (Este supuesto recibe a menudo el nombre de «el principio del caos molecular».) El supuesto puede interpretarse también en un sentido puramente estadístico —produciendo las mismas dificultades que antes— o en un sentido propensivista. En este último sentido, la hipótesis significa que (i) las condiciones experimentales bajo control no fijan absolutamente las condiciones iniciales, sino que les dejan cierto juego, y que (ii) cada una de las posibilidades que se han dejado así abiertas a las condiciones iniciales se realizarán con una cierta propensión o probabilidad (que puede calcularse a veces con la ayuda de consideraciones de simetría). Uno de los méritos del argumento de Landé es que muestra que esas consideraciones más satisfactorias han entrado sin ser advertidas, y también que deberían haber sido absolutamente excluidas por aquellos que desean retener un sistema determinista.

b) Para poder mostrar con mayor claridad aún las dificultades en las que se encuentra el determinista, supongamos que la proporción de las bolas *d* con respecto a las bolas *i* no era de 50 : 50, sino, digamos, de 40 : 60. En este caso, es razonable suponer que un diminuto desplazamiento de la cuchilla hacia la izquierda cambiará la proporción en favor de las bolas *d.* Quizá podamos conseguir, a consecuencia del desplazamiento, una proporción de 52 : 48 ó 50 : 50, y un desplazamiento ligeramente mayor puede incluso dar la mayoría a las bolas *d.*

Se concederá que experimentos con resultados como ésos pueden hacerse si se da por sentado que podemos obtener frecuencias de Landé estables; es decir, todos estamos dispuestos a *predecir* que pequeños ajustes en la posición de la cuchilla llevan a resultados como los indicados. Pero, para el determinista, las predicciones de este tipo tienen que ser imposibles, o milagrosas, puesto que tendrían que suponer una «armonía pre-establecida» en las condiciones iniciales; el determinista no puede, como hemos visto, explicarlas.

c) El argumento de Landé puede usarse también para criticar la doctrina de que las consideraciones de probabilidades sólo entran en la ciencia si *nuestros conocimientos* son insuficientes para permitirnos hacer predicciones con certidumbre.

Desde un punto de vista determinista, esta doctrina es de una importancia fundamental: es la única alternativa a la teoría de la distribución estadística irreductible y milagrosa de las condiciones iniciales que acabamos de criticar. Obviamente, es la única doctrina que, desde un punto de vista determinista, puede explicar enunciados de probabilidad *singular.* Pero parece que la sostienen muchos que no profesan opiniones deterministas [14].

Para ver la debilidad e incluso la irrelevancia de esta doctrina, supongamos una vez más que nos encontramos con una situación como la descrita por Landé, con bolas que caen sobre una cuchilla de acero y una proporción de 50 : 50 de bolas *d* y bolas *i.* Supongamos también que tenemos una cuchilla óptica con cuya ayuda podemos *saber con certeza* de cada bola si va a ser derecha o izquierda. Esto, indudablemente, hace innecesario invocar probabilidades respecto a la predicción de cada bola individual. Pero esto no afecta a nuestro problema de ninguna manera. Podemos suponer [15] que las bolas caen a la derecha o a la izquierda de la cuchilla de acero exactamente como antes, con la misma proporción de 50 : 50 y con las mismas fluctuaciones estadísticas; y el problema de explicar esos resultados estadísticos y el de explicar nuestra capacidad de predecir que las secuencias futuras llevarán a resultados similares (siempre que las condiciones permanezcan iguales) continúan siendo exactamente los mismos que antes, a pesar de que *ahora sabemos cada resultado individual por anticipado.*

Pero ¿no nos permite nuestro conocimiento anticipado de las bolas *d* y de las bolas *i* cambiar sus proporciones? Podemos suponer que las bolas pasan por el tubo de Landé con suficiente lentitud y suficientemente separadas entre sí como para observarlas

---

[14] Por ejemplo, Pauli lo mantiene (véase *Teoría cuántica y el cisma en Física,* tomo III de este *Post Scriptum,* sección 5). Lo mantienen, por regla general, todos aquellos que apoyan una interpretación subjetiva o inductiva de la probabilidad. Hasta qué punto la razón de apoyar esas interpretaciones surge de una creencia inconsciente en el determinismo (como sospecho yo) es, naturalmente, difícil de decidir.

[15] Si las bolas son suficientemente pesadas, no las afectará nuestra cuchilla óptica.

con la cuchilla óptica y sacar cada bola *d* con la mano (y meterla en una caja, por ejemplo). Como consecuencia, sólo obtendremos bolas *i*, en lugar de una proporción de 50 : 50. Así, sobre la base de nuestro conocimiento preciso, podemos *controlar* los resultados estadísticos a nuestro antojo.

Ciertamente, este argumento es correcto. Pero seguiremos encontrando que la proporción de las bolas *i* y de las bolas que hemos puesto en la caja sigue siendo 50 : 50, como antes; y el problema de explicar esta proporción y las fluctuaciones estadísticas sigue siendo el mismo: no hemos hecho más que desplazarlo.

La proporción de 50 : 50, ahora ya debe estar claro, depende de las condiciones objetivas experimentales, y no tiene absolutamente nada que ver con nuestro conocimiento, o la falta de él. En la medida en que cambiamos las condiciones experimentales, sustituyendo las bolas *d* por bolas en-caja, hubo un cambio en los resultados; y en la medida en que no cambiamos las condiciones, dejando el tubo y la cuchilla en sus sitios, no hubo cambio.

30. LA CUCHILLA DE LANDE Y LA INTERPRETACION PROPENSIVISTA

Hemos visto que cuando desplazamos la cuchilla cambian las frecuencias [*cf. b)* en la sección anterior]. La tarea de la teoría será explicar este hecho de una manera sencilla, mostrando por qué podemos predecir esos y otros cambios similares (como de hecho hacemos).

Cualquier cambio en la posición de la cuchilla cambia las posibilidades inherentes a la situación experimental y a sus condiciones de simetría. Más precisamente, cambia la medida de dichas posibilidades: un desplazamiento a la izquierda aumenta la posibilidad de obtener bolas *d*. Llamando a las medidas de las posibilidades probabilidades objetivas o *propensiones*, no estoy haciendo nada más que usar otra palabra; pero lo hago para llamar la atención sobre el hecho de que esas «posibilidades» se consideran ahora *magnitudes físicas* que, como las fuerzas, pueden interactuar y combinarse, y pueden, por tanto, considerarse, a pesar del término «posibilidad», como *físicamente reales:* no son simplemente posibilidades lógicas, sino que son *posibilidades físicas.*

Las propensiones pueden interpretarse como probabilidades

objetivas, singulares. Son singulares en la medida en que son inherentes a la situación experimental que se supone igual en todos los experimentos. (Así obtenemos independencia, o inmunidad contra efectos secundarios, para los elementos de la secuencia.) Y las propensiones se manifiestan así de forma bernouilliana, en las frecuencias de la secuencias que son repeticiones de la situación experimental.

Si la situación experimental es tal que siempre obtenemos el mismo resultado —por ejemplo, sólo bolas *d*—, entonces puede ser de tipo determinista *prima facie;* si es tal que obtenemos frecuencias relativas que no son iguales a 1 ni a 0, entonces será de tipo probabilista. En cada uno de los casos podemos decir que la situación experimental determina las *probabilidades* de cada resultado individual del experimento o las *propensiones* a obtener ciertos resultados.

Como las condiciones son condiciones físicas objetivas, las propensiones o probabilidades son también objetivas. Deberían considerarse propiedades, no del sistema bajo investigación (la bola, el electrón, o lo que sea), sino como propiedades de toda la situación experimental (que incluye, naturalmente, la bola, el electrón; es decir, el sistema en investigación).

Así pues, sugiero que admitamos que pueden existir las propensiones —igual que las fuerzas, u otras entidades físicas abstractas u ocultas, que se han introducido para explicar lo conocido por medio de lo desconocido. Son, como las fuerzas, el resultado (o dependen) de ciertas *relaciones* entre otras entidades físicas; digamos, entre cuerpos físicos, o entre entidades más abstractas, tales como corrientes, de campos, o quizá incluso entre otras probabilidades: una vez que admitimos esas entidades físicas abstractas, pero objetivas (que pueden ser capaces de cambiar continuamente) en nuestra teoría física, no hay razón para que no las dejemos interactuar, o ser relacionadas por leyes que, en determinadas situaciones, hacen la propensión en un lugar dependiente de las propensiones de su entorno.

Mi vieja idea, desarrollada en *La lógica de la investigación científica* (por ejemplo, en la sección 57), era que los resultados estadísticos (como los tratados por Landé) debían ser explicados por hipótesis estadísticas que, a su vez, pudieran ser inspiradas por consideraciones de simetría, pero no pudieran deducirse de ellas.

Esta concepción fue criticada por Einstein (en dos cartas) y

también por Jordan [16]. Ambos habrían tenido razón si hubieran afirmado solamente que yo estaba equivocado; pero los dos estaban equivocados en su argumento de que los resultados estadísticos pueden obtenerse a partir de supuestos deterministas clásicos. Las premisas probabilistas son efectivamente indispensables para las conclusiones estadísticas, aunque esas premisas no tienen que ser estadísticas, pero pueden ser hipótesis sobre propensiones, y puesto que las propensiones son medidas de posibilidades pueden, en ciertos casos, derivarse, con toda validez, de consideraciones de simetría (éste es el caso en el ejemplo de Einstein), o del hecho de que ciertas posibilidades tienen la medida cero (éste es el caso del ejemplo de Jordan).

31. CONCLUSION

Aunque creo que mi refutación del determinismo «científico» es necesaria para preparar la base para la plena comprensión de la probabilidad en física, mi propia refutación (distinta de la de Landé) no utiliza en ningún momento la teoría de la probabilidad; ni tampoco apelo a la teoría cuántica. El «libre albedrío» tampoco se menciona más que de paso. [Véase, sin embargo, el «Epílogo».] Mi argumento es válido para toda teoría física, por más que ésta parezca fuertemente determinista.

Respecto a su aplicación a cuestiones humanas y a los problemas de la ética y de la responsabilidad, sólo se han dado unas pocas pistas (en las secciones 15, 16, 23 y 24). Todos los sucesos singulares de este mundo son únicos y, si se los considera bajo el aspecto de su unicidad, pueden describirse como indeterminados o «libres». Para algunos sucesos, esta forma de describirlos quizá sea forzada. Pero, cuando se trata de personalidades humanas y de sus acciones, ése puede ser el aspecto más importante para nosotros. Eso es lo que ocurre claramente cuando estamos personalmente interesados en las personas de las que se trata.

En la medida en que el hombre es un predictor, mis resultados sobre las máquinas predictoras son, creo yo, aplicables *a fortiori* al hombre y a la sociedad humanas.

---

[16] Una de las cartas de Einstein se ha editado como Apéndice *XII a *La lógica de la investigación científica [L. Sc. D.]*; el pasaje al que me refiero es el último párrafo de la carta. La crítica de P. Jordan se encuentra en su *Anschauliche Quantentheorie*, 1936, p. 282.

«Conócete a ti mismo» —es decir, conoce tus limitaciones— es un ideal que, ahora lo sabemos, es lógicamente irrealizable. Como somos calculadoras, no podemos conocernos plenamente, ni siquiera podemos conocer todas nuestras limitaciones; al menos no las limitaciones a nuestro conocimiento.

Pero, ciertamente, no pretendo sugerir que el paralelismo entre nosotros y la máquina de predecir va lejos. Pienso que el hombre no es sólo un predictor, sino más. Incluso en lo que respecta a nuestras actividades puramente intelectuales, tenemos esperanzas, temores, intereses y problemas. No somos simplemente calculadoras ni somos principalmente calculadoras; y en la medida en que lo somos, somos unas calculadoras miserablemente malas. Cualquier máquina corriente de sumar es superior a la mayoría de nosotros. En realidad, no construiríamos tablas de multiplicar y sistemas de aritmética si nuestro cerebro pudiera calcular. Construimos métodos de calcular con lápiz y papel, y cerebros electrónicos, simplemente porque no tenemos suficiente cerebro nosotros mismos.

Así, no somos principalmente calculadoras, pero somos constructores de ellas. Las hacemos porque estamos interesados en problemas cuyas soluciones están por encima de nuestra limitada capacidad de calcular, y, aún más, porque nos fascinan los nuevos problemas que nos plantea la construcción de calculadoras. Nuestro impulso intelectual fundamental es buscar dificultades, o incluso inventar dificultades para superarlas.

Una calculadora, quizá, pueda crear teoremas matemáticos. Quizá pueda distinguir entre pruebas y no-pruebas y, debido a ello, entre ciertos teoremas y no-teoremas. Pero no, no distinguirá las pruebas difíciles e ingeniosas y los teoremas interesantes de los que son aburridos y carecen de interés. Por tanto, una calculadora «sabrá» muchísimas cosas —demasiadas— que no tienen interés. El saber de una calculadora, por muy sistemático que sea, es como un mar de trivialidades en el que puede haber suspendidas unas cuantas partículas de oro, de información valiosa. (Coger esas partículas puede ser tan difícil como tratar de obtenerlas sin la calculadora, y más aburrido.) Sólo el hombre, con sus problemas, puede dar significado a la capacidad absurda de la calculadora de producir verdades.

Para expresar mi argumento de una forma más seria, es función de toda teoría dividir todos los enunciados referidos al tema en cuestión en tres clases distintas: aquellos de los que la teoría

afirma que son verdaderos, aquellos de los que afirma que son falsos y aquellos sobre los que no hace afirmación alguna. Por esta razón es por lo que una teoría inconsistente es inútil; porque no logra esta división, sino que afirma todo enunciado (y, por tanto, también la negación de todo enunciado). Una teoría inconsistente es inútil porque afirma demasiadas cosas.

Ahora bien, una buena calculadora (es decir, una consistente) no es ciertamente inútil, puesto que puede realizar dicha clasificación. Sin embargo, sigue afirmando demasiadas cosas. Si se le hace derivar todas las consecuencias de cualquier teoría, una tras otra, automáticamente, seguirá sin tener un método de escoger las consecuencias interesantes o importantes, ni tendrá un método de asegurarse de que produce siquiera una de ellas dentro de un intervalo de tiempo determinado. Para cada enunciado de utilidad moderada, tal como «2 + 1 = 3», contendrá también una secuencia infinita de enunciados «2 + 1 ≠ 4», «2 + 1 ≠ 5»... En la secuencia infinita de los enunciados en su orden de producción, la probabilidad de dar con uno interesante será (con cualquier criterio razonable) cero.

Sólo el cerebro humano [* o quizá debería haber dicho la mente humana] puede crear intereses, propósitos, problemas y objetivos, incluso dentro del terreno relativamente estrecho de sus actividades intelectuales.

Otro argumento sería éste: aprendemos mediante errores, y eso significa que, cuando llegamos a incoherencias, volvemos atrás y volvemos a formular nuestros supuestos. Al aplicar este método llegamos al extremo de reconsiderar los supuestos, incluso los de naturaleza lógica, si ello es necesario. (Esto ocurrió en el caso de las paradojas lógicas.) Es difícilmente concebible que una máquina pueda hacer lo mismo. Si sus incautos creadores la equipan con incoherencias, entonces, a su debido tiempo, deducirá todos los enunciados que pueda formular (y sus negaciones). Quizá podamos equiparla con un aparato que la avise, en caso de que deduzca que «0 = 1», y le haga abandonar algunos de sus supuestos. Pero difícilmente podremos construir una máquina que pueda criticar y reajustar sus propios métodos de deducción o sus propios métodos de crítica.

La conclusión general de nuestras consideraciones parece ser la restauración de la concepción ingenua del mundo, descrita como «la concepción según el sentido común» en la sección 1:

la perspectiva según la cual hay sucesos que pueden predecirse o que están «determinados» y otros sucesos que no pueden predecirse y que no están «determinados».

Pero nuestras consideraciones sugieren, incluso, algo parecido a una reconciliación entre esta perspectiva y la otra —la «más sofisticada»—, según la cual, en general, es sólo la falta de conocimiento lo que nos lleva a creer que los sucesos son impredictibles.

Esta reconciliación puede producirse si nos percatamos de que la existencia de conocimiento en el mundo físico —o, más bien, de sucesos físicos que pueden interpretarse como representantes del conocimiento, o resultantes del conocimiento— da lugar al tipo de indeterminismo que hemos tratado aquí. El saber puede conquistar nuevos problemas. Pero al hacerlo creará nuevos problemas que no podrá resolver; al menos, no inmediatamente. Porque no puede conocer de antemano sus propias conquistas futuras.

ADDENDA

EL INDETERMINISMO NO BASTA: UN EPILOGO [1]

Lo que me ocupa aquí es la libertad humana. Con libertad humana me refiero a lo que se ha llamado habitualmente «libre voluntad». Sin embargo, evitaré el término «voluntad»; no vaya a ser que el discutirlo nos desvíe a estériles problemas de terminología. Por razones similares, no trataré explícitamente de la libertad moral, a pesar de que es el tipo de libertad humana por la que los filósofos se han interesado con mayor frecuencia. En cambio, empezaré hablando de la libertad de crear obras de arte o teorías explicativas en ciencia. La libertad moral tiene gran importancia, pero, una vez más, el discutir sobre ella puede desviarnos hacia el problema de la responsabilidad moral o incluso hacia el problema de la recompensa y el castigo. Espero, sin embargo, simplificar mi exposición evitando cualquier discusión directa de cuestiones morales y ciñéndome a las de la libertad de crear, y la libertad de evaluar razones o argumentos, en pro o en contra de enunciados de hecho, o teorías científicas. Presumiblemente, si disfrutamos de *este* tipo de libertad, es posible que disfrutemos también de la libertad de crear, de razonar y de deducir en los aspectos morales, y de disfrutar de la responsabilidad que acompaña a la creación y a la elección. Mientras que, si *no* tenemos, como mínimo, la libertad de razonar y de argüir sobre cuestiones de hecho, difícilmente podemos tener libertad moral alguna.

El título de este ensayo-epílogo, «El indeterminismo no basta», quiere indicar que una física indeterminista —que es lo que el núcleo central de este libro trata de defender— no es suficiente

---

[1] [Desde que se escribió el *Post Scriptum*, hace más de veinticinco años, Popper ha publicado varios ensayos relacionados con las cuestiones del determinismo y la libertad humana. Uno de ellos, la conferencia en memoria de Arthur Holly Compton, «Sobre nubes y relojes. Aproximación al problema de la racionalidad y libertad del hombre», se ha publicado (como capítulo 6 de *Conocimiento objetivo* [*Obj. Kn*] y es fácil de obtener. Un segundo ensayo, «Indeterminism Is Not Enough» —que se publicó, por primera vez, en *Encounter*, 40, abril 1973, páginas 20-26—, no es tan fácil de conseguir y constituye una importante adición al argumento de este libro. Por tanto, se publica aquí como una especie de Epílogo a este libro. *Ed.*]

por sí misma para crear espacio para la libertad humana: no es suficiente hacer inteligible la libertad humana. Para crear ese espacio, afirmo, necesitamos más. Necesitamos además, como mínimo, la *apertura causal* de lo que voy a llamar el «Mundo 1» hacia el «Mundo 2», y la *apertura causal* del «Mundo 2» hacia el «Mundo 3», y viceversa. Así pues, empezaré por explicar lo que llamo «Mundo 1», «Mundo 2» y «Mundo 3» [2].

Los «Mundos 1, 2 y 3»

Como «Mundo 1» me refiero a lo que puede llamarse el mundo de la física: de las rocas, los árboles y los campos físicos de fuerzas. También incluyo aquí los mundos de la química y la biología. Con «Mundo 2» me refiero al mundo psicológico. Lo estudian los estudiosos de la mente humana, pero también los de la mente animal. Es el mundo de los sentimientos de temor y esperanza, de las disposiciones a actuar y de todo tipo de experiencias subjetivas, incluidas las subconscientes e inconscientes. Así, los términos «Mundo 1» y «Mundo 2» quedan fácilmente explicados. La explicación de lo que llamo «Mundo 3» es un poco más difícil.

Con «Mundo 3» me refiero al mundo de los productos de la mente humana. Aunque incluyo las obras de arte en el «Mundo 3» y también los valores éticos y las instituciones sociales (y así, podríamos decir, las sociedades), me limitaré, principalmente, al mundo de las bibliotecas científicas, a los libros, a los problemas científicos, y a las teorías, incluidas las erróneas.

Los libros, periódicos y bibliotecas pertenecen tanto al «Mundo 1» como al «Mundo 3». Son objetos físicos y, como tales, pertenecen al «Mundo 1»: están sometidos a las restricciones físicas de las leyes físicas del «Mundo 1». Por ejemplo, aunque dos ejemplares del mismo libro sean totalmente iguales físicamente, no pueden ocupar la misma parte del espacio físico; por tanto, son dos objetos diferentes del «Mundo 1». Pero no sólo pertenecen al «Mundo 1»: también pertenecen al «Mundo 3». *Dos* ejemplares muy parecidos del mismo libro son *diferentes* como objetos del «Mundo 1»; pero si el *contenido* de dos libros físicamente iguales

[2] [Para unas explicaciones más detalladas de los Mundos 1, 2 y 3, véase K. R. Popper: *Conocimiento objetivo [Obj. Kn.]*, capítulos 3 y 4, y *El yo y su cerebro* —con Sir John Eccles— [*S. & I. B.*, 1977], capítulo P 2. *Ed.*]

(o distintos) es el mismo, entonces, como objetos del «Mundo 3», los dos libros son idénticos: son dos copias de *un* objeto del «Mundo 3». Aún más, este objeto del «Mundo 3» está sujeto a las restricciones y valoraciones del «Mundo 3»; por ejemplo, puede ser examinado por su coherencia lógica, y valorado por su contenido informativo.

El *contenido* de un libro, o de una teoría, es algo abstracto. Todos los cuerpos físicos *concretos*, tales como las rocas, los árboles y los cuerpos animales y humanos, pertenecen al «Mundo 1»; y todos los estados psicológicos, sean conscientes o subconcientes, pertenecen al «Mundo 2». Pero las cosas *abstractas*, como los problemas, teorías y argumentos, incluidos los erróneos, pertenecen al «Mundo 3». (También los argumentos y teorías inconsistentes. Esto no hace, naturalmente, inconsistente al «Mundo 3», porque el «Mundo 3» no es una teoría ni una aserción ni un argumento: es una clase de cosas, el universo del discurso.) Además, a menos que decidamos introducir, para las obras de arte, un nuevo término, como «Mundo 4», una obra como *Hamlet* y una sinfonía como la *Inacabada* de Schubert pertenecen también al «Mundo 3»; y, de la misma forma que un ejemplar determinado de un libro pertenece *tanto* al «Mundo 1» *como* al «Mundo 3», así las realizaciones determinadas de *Hamlet* y de la sinfonía *Inacabada* de Schubert pertenecen tanto al «Mundo 1» como al «Mundo 3». Pertenecen al «Mundo 1» en cuanto que consisten en sucesos físicos complejos; pero pertenecen al «Mundo 3» en cuanto que tienen un contenido, un mensaje, un significado.

Los términos «Mundo 1», «Mundo 2» y «Mundo 3» se han escogido conscientemente por su falta de color y por la arbitrariedad. Pero hay una razón histórica para llamarlos 1, 2 y 3: parece que el mundo físico existió antes que el mundo de los sentimientos animales; y yo conjeturo que el «Mundo 3» no comienza hasta la evolución de un lenguaje humano específico. Consideraré que el mundo del *conocimiento humano formulado lingüísticamente* es el más característico del «Mundo 3». Es el mundo de los problemas, las teorías y los argumentos; e incluiré también aquellos problemas, teorías y argumentos que no han sido formulados lingüísticamente todavía. También supondré que el «Mundo 3» tiene historia: que ciertos problemas, teorías y argumentos fueron descubiertos, o quizá refutados, en determinadas fechas, mientras que otros, en esas mismas fechas, no habían sido descubiertos, o refutados, aún.

La realidad de los tres Mundos

Es, creo yo, de buen sentido común aceptar la realidad o existencia del «Mundo 1» de los cuerpos físicos. Como muestra la famosa refutación de Berkeley por el doctor Johnson, puede decirse de un cuerpo físico, tal como una roca, que existe, porque puede dársele una patada, y si das una patada suficientemente fuerte a una roca, notarás que te la devuelve. Siguiendo a Alfred Landé, propongo que se diga que algo existe, o es real, si, y sólo si, puede recibir una patada y puede, en principio, devolverla; para expresarlo en términos algo más generales, propongo que se diga que algo existe, o es real, si, y sólo si, puede interactuar con los miembros del «Mundo 1», con cuerpos duros, físicos.

Así pues, el «Mundo 1», o el mundo físico, puede considerarse como el ejemplo típico de la realidad o la existencia. Sin embargo, yo creo que las cuestiones de terminología, o del uso o significado de las palabras, son insignificantes. Así, considero que el uso de palabras como «real» o «existente» no es muy importante; especialmente si se compara con cuestiones sobre la *verdad* de las afirmaciones o proposiciones teóricas.

La proposición cuya verdad quiere defender, y que me parece a mí que va un poco más lejos del sentido común, es que no sólo son reales el físico «Mundo 1», y el psicológico «Mundo 2», sino que también lo es el abstracto «Mundo 3», *real* exactamente en el mismo sentido en el que es real el «Mundo 1» físico, de rocas y árboles: los objetos de los «Mundos 2 y 3» pueden darse patadas mutuamente, de la misma forma que los objetos del «Mundo 1»; y también pueden recibirlas.

La realidad de los «Mundos 1 y 2»

Aunque propongo, con el doctor Johnson y con Alfred Landé y con otros realistas del sentido común, considerar al «Mundo 1» como el patrón mismo de realidad, yo no soy un monista, sino un pluralista [3]. Un inmaterialismo o fenomenalismo monista que negaba la existencia del «Mundo 1» y admitía sólo la existencia de las experiencias, es decir, sólo el «Mundo 2», estuvo bastante de moda hasta hace relativamente poco. Actualmente, está mu-

[3] Véase *El yo y su cerebro* [*S. & I. B.*], capítulos P 3 y P 5.

cho más de moda la opinión contraria. Es decir, la opinión de que *sólo* existe el «Mundo 1». Esta teoría se llama materialismo monista o fisicalismo o conductismo filosófico. Más recientemente, esta teoría se ha llamado también la «teoría de la identidad» porque afirma que las experiencias mentales son, en realidad, idénticas a los procesos del cerebro.

Las diversas formas del monismo van a sustituirse aquí por un pluralismo: la tesis de los tres mundos. Este pluralismo puede apoyarse en dos líneas muy diferentes de argumentación. Primero, para demostrar la realidad del «Mundo 2», puede apelarse al sentido común y al fracaso de los fisicalistas a la hora de producir argumentos efectivos en contra de la opinión del sentido común de que un dolor de muelas puede ser ciertamente muy real.

Sin embargo, mi segundo argumento, el principal, sigue un curso muy diferente. Parte de la afirmación de que los objetos del «Mundo 3», tales como las teorías, interactúan de hecho, con fuerza, con el «Mundo 1», físico. Los ejemplos más simples son los modos en que hacemos cambios en el «Mundo 1» cuando construimos, digamos, reactores nucleares o bombas atómicas o rascacielos o aeropuertos, de acuerdo con planes y teorías del «Mundo 3» que son, con frecuencia, sumamente abstractos.

Mi principal argumento en favor de la existencia del «Mundo 2» de las experiencias subjetivas es que normalmente tenemos que captar o entender una teoría del «Mundo 3» antes de poder usarla para actuar sobre el «Mundo 1», pero captar y entender una teoría es un asunto mental, un proceso del «Mundo 2»: el «Mundo 3» suele interactuar con el «Mundo 1» a través del «Mundo 2» mental. Un ejemplo de ello es la planificación, la construcción y el uso de excavadoras para construir un aeropuerto. Hay, al principio, una interacción entre la planificación por la mente humana, en el «Mundo 2», y las restricciones internas tanto del «Mundo 1» como del «Mundo 2» que limitan la planificación de la maquinaria. En segundo lugar, tenemos una interacción entre el «Mundo 2» y el «Mundo 1» del cerebro humano que, a su vez, actúa sobre nuestros miembros que manejan las excavadoras.

La efectividad de este argumento depende claramente del «Mundo 3». Si existe el «Mundo 3» y es, al menos en parte, autónomo, y si, además, los planes del «Mundo 3» afectan realmente al «Mundo 1», entonces me parece inescapable que también exista un «Mundo 2». De este modo, mi principal argumento en favor

de la existencia del «Mundo 2» nos ha vuelto a llevar al problema de si existe el «Mundo 3»; y, después de eso, al problema de si el «Mundo 3» es, en parte, autónomo.

La realidad y parcial autonomía del «Mundo 3»

El lenguaje humano y el pensamiento humano evolucionan juntos, en mutua interacción. El lenguaje humano, no hay duda, expresa los procesos del pensamiento humano; es decir, objetos del «Mundo 2». Pero existe una gran diferencia cuando esos objetos del «Mundo 2», subjetivos, son formulados en un lenguaje humano objetivo: hay un poderoso efecto de realimentación entre el lenguaje humano y la mente humana.

Este ocurre sobre todo porque un pensamiento, una vez formulado en lenguaje, se convierte en un *objetivo* fuera de nosotros. Ese objeto puede ser criticado inter-subjetivamente, criticado por otros y por nosotros mismos. La crítica objetiva o inter-subjetiva en este sentido surge sólo con el lenguaje humano; y con él surge el «Mundo 3», humano; el mundo de las normas objetivas y de los contenidos de nuestros procesos de pensamiento subjetivo.

Así pues, es muy diferente *pensar* simplemente un pensamiento o *formularlo* en un lenguaje (o, aún mejor, ponerlo por escrito o hacerlo imprimir). Mientras no hagamos más que pensar el pensamiento, no puede ser criticado objetivamente. Es parte de nosotros mismos. Para ser criticable tiene que ser formulado en un lenguaje humano y convertirse en un objeto: un objeto del «Mundo 3». Los pensamientos formulados lingüísticamente pertenecen al «Mundo 3». Pueden ser criticados *desde un punto de vista lógico,* por ejemplo, demostrando que tienen ciertas consecuencias lógicas inoportunas o incluso absurdas. Sólo los *contenidos de pensamiento* que pertenecen al «Mundo 3» pueden estar en relaciones lógicas, tales como equivalencia, deducibilidad o contradicción.

Así podemos distinguir claramente entre los *procesos de pensamiento* subjetivo, que pertenecen al «Mundo 2», y los *contenidos* objetivos de los pensamientos, los contenidos en sí mismos, por así decirlo, que constituyen el «Mundo 3».

Para dejar más claro el argumento, tomemos a dos matemáticos que, equivocándose, llegan ambos a un teorema falso; por ejemplo, al teorema de que $5 + 7 = 13$. Sus procesos de pensamiento,

que pertenecen al «Mundo 2», pueden ser iguales o enteramente diferentes. Pero el *contenido* de sus pensamientos, que pertenecen al «Mundo 3», es uno y el mismo, y puede ser criticado. Los dos matemáticos pueden recibir la patada de la estructura lógica del «Mundo 3» que muestra que su supuesto teorema *contradice* el enunciado objetivamente verdadero «5 + 7 = 12», y que, por tanto, debe de ser objetivamente *falso*. Los dos matemáticos reciben una patada, no de otra gente, sino de las propias leyes de la aritmética.

La mayoría de las personas son dualistas: forma parte del sentido común creer en la existencia de los «Mundos 1 y 2». Pero no es tan fácil para la mayoría de la gente aceptar la existencia del «Mundo 3». Admitirán, por supuesto, que existe una parte muy especial del «Mundo 1» que consiste en libros impresos, o de sonidos de lenguaje acústicos, y admitirán los procesos del cerebro y los procesos del pensamiento subjetivo. Pero afirmarán que lo que distingue a los libros de otros cuerpos físicos, como los árboles, o al lenguaje humano de otros ruidos, como el aullido de los lobos, es *sólo* el hecho de que nos ayudan a tener ciertos tipos especiales de experiencias propias del «Mundo 2», a saber, procesos de pensamiento de un tipo especial (que quizá corran paralelos a los procesos del cerebro) que están correlacionados precisamente con esos libros o esos ruidos lingüísticos.

Considero que esa forma de ver las cosas es totalmente insuficiente. Trataré de mostrar que tenemos que admitir la existencia de una parte autónoma del «Mundo 3»; una parte que consiste en *los contenidos de pensamiento* objetivos que son *independientes,* y claramente diferenciables de los procesos de pensamiento subjetivos o personales por medio de los cuales son captados y en cuya captación influyen causalmente. Afirmo así que existen objetos autónomos del «Mundo 3» que no han tomado aún una forma propia del «Mundo 1» ni del «Mundo 2», pero que, no obstante, interactúan con nuestros procesos de pensamiento. De hecho, influyen decisivamente en nuestros procesos de pensamiento.

Tomemos un ejemplo de la aritmética elemental. La sucesión infinita de los números naturales 0, 1, 2, 3, 4, 5, 6, etc., es una invención humana, un producto de la mente humana. Como tal, puede decirse que *no* es autónoma, sino que depende de los procesos del «Mundo 2». Pero tomemos ahora los números pares a los números primos. Esos no han sido inventados por nosotros,

sino *descubiertos* o encontrados. *Descubrimos* que la sucesión de los números naturales consta de números pares y de números impares y, sea lo que fuere lo que pensemos de ello, ningún proceso de pensamiento puede alterar este hecho del «Mundo 3». La sucesión de los números naturales es un resultado de que hayamos aprendido a contar; es decir, es un invento dentro del lenguaje humano. Pero tiene sus leyes o restricciones o regularidades internas inalterables que son las *consecuencias no buscadas* de la sucesión, creada por el hombre, de los números naturales; es decir, las consecuencias no buscadas de un producto de la mente humana.

Lo mismo puede decirse de los números primos. Se descubrió que cuanto más se eleva uno en la sucesión de los números naturales (por ejemplo, primero en los números entre 100 y 200 y luego de 1.100 a 1.200), la ocurrencia de números primos se hace más rara: ésta es una *propiedad autónoma* del «Mundo 3». Ahora bien, este descubrimiento nos lleva a un nuevo problema autónomo del «Mundo 3»; un problema que se descubre, que simplemente está ahí, igual que los propios números primos. Se trata de este interesante problema: si avanzamos hacia números cada vez mayores, hasta 10 millones, por ejemplo, ¿acaban desapareciendo los números primos, o hay *siempre* nuevos números primos, aunque vayan siendo cada vez más escasos?, o, en la terminología de Euclides, ¿existe un número primo máximo o es la sucesión de los números primos infinita, igual que la sucesión de los números naturales?

Este es un problema objetivo y autónomo: o bien existe un número primo máximo, o la secuencia de los números primos continúa hasta el infinito. Euclides, que incluso quizá fue quien descubrió el problema, lo resolvió. Demostró que la hipótesis de que existe un número primo máximo lleva a un absurdo.

De esta forma dio una prueba, una prueba indirecta, del hecho objetivo de que no existe un número primo máximo, sino que siempre existe uno mayor: la sucesión de los números primos es infinita, igual que la de los números naturales. Y este hecho es un hecho objetivo, autónomo, del «Mundo 3». Es un teorema del «Mundo 3». Podemos descubrirlo, podemos demostrarlo; pero nada podemos hacer para cambiarlo.

El descubrimiento de los números primos ha llevado a muchos problemas difíciles, algunos de los cuales han sido resueltos, mientras que otros muchos siguen abiertos. Esos problemas son

descubiertos por nosotros en el nuevo campo que hemos creado, en la sucesión de los números naturales. Se descubre que están ahí independientemente de que alguien haya pensado sobre ellos anteriormente, o no. Así, tenemos construcciones en matemáticas que son productos de la mente humana y problemas y teorías que son las consecuencias objetivas, unas consecuencias sobre las que quizá nadie había reflexionado, de esas construcciones. Esto muestra que el mundo de la matemática contiene una parte autónoma: una parte autónoma del «Mundo 3».

Mi siguiente argumento es que esta parte autónoma del «Mundo 3» es *«real»* en el sentido de que puede interactuar con el «Mundo 2» y también a través del «Mundo 2» con el «Mundo 1». Si algunos hombres, o muchos, buscan una solución para un problema matemático aún no resuelto, entonces todos ellos están, posiblemente de muchas maneras diferentes, influidos por ese problema. El éxito de sus intentos de resolverlo dependerá, al menos una parte, de la existencia o inexistencia, en el «Mundo 3», de una solución al problema, y, en parte, de si sus procesos de pensamiento les llevan o no a contenidos de pensamiento objetivamente verdaderos. Esto muestra que los objetos autónomos del «Mundo 3» ejercen una fuerte influencia causal sobre los procesos del «Mundo 2». Y si un problema, recientemente descubierto, del «Mundo 3», con o sin solución, se publica, entonces la influencia causal se extiende incluso al «Mundo 1», al contribuir a poner en movimiento los dedos de los tipógrafos e incluso las ruedas de las máquinas impresoras.

Por razones tan sencillas como ésas, yo mantengo no sólo que el «Mundo 3» es parcialmente autónomo, sino, además, que su parte autónoma es real, puesto que puede actuar sobre el «Mundo 1», al menos a través del «Mundo 2». La situación es esencialmente la misma con cada descubrimiento científico y cada invento técnico. En todos esos casos, los problemas del «Mundo 3» y sus teorías juegan un papel primordial. Los problemas pueden descubrirse y, aunque las teorías (que son, digamos, sobre el «Mundo 1») sean productos de la mente humana, *no son simplemente* construcciones nuestras; porque su verdad o falsedad dependen enteramente de su relación con el «Mundo 1», una relación que, en todos los casos importantes, no podemos alterar. Su verdad o falsedad dependen de la estructura interna del «Mundo 3» (especialmente lenguaje) y del «Mundo 1», el último de los cuales, como he sugerido, es el patrón mismo de la realidad.

La situación humana y el mundo natural

El origen de la vida puede ser, por lo que sabemos en la actualidad, una ocurrencia única en el universo. No podemos explicarlo y se acerca mucho a lo que David Hume hubiera llamado, a regañadientes, un milagro. La emergencia de un «Mundo 2» de conciencia animal, de sentimientos de alegría y de dolor, parece ser el segundo milagro.

Parece razonable considerar la emergencia de la conciencia y, anteriormente, la de la vida como dos acontecimientos relativamente recientes en la evolución del universo; como acontecimientos que, como el principio del universo, están actualmente, y quizá lo estarán siempre, más allá de nuestra comprensión científica. Este modesto enfoque admite libremente la existencia de problemas abiertos y, por tanto, no cierra el camino para descubrir más cosas sobre ellos, sobre su carácter y, quizá, incluso sobre los caminos hacia una posible solución o, al menos, hacia una solución parcial.

El tercer gran milagro es la emergencia del cerebro humano, y de la mente y la razón humanas. Este tercer milagro puede estar menos alejado que los otros de una explicación, al menos en términos evolutivos. El hombre es un animal. Parece estar mucho más cerca de otros animales que éstos de la materia inanimada. Pero esto no significa desestimar el abismo que separa el cerebro humano del cerebro animal, y el lenguaje humano de todos los demás lenguajes animales, de las disposiciones que poseen la mayor parte de los animales superiores para *expresar* sus estados íntimos para *comunicarse* con otros animales.

El hombre ha creado el lenguaje humano, con su función *descriptiva* y el valor de la verdad, y con su función *argumentativa* y el valor de la validez de los argumentos, trascendiendo así los lenguajes animales con sus funciones meramente expresiva y comunicativa [4]. Con él, el hombre ha creado el «Mundo 3» objetivo, algo para lo que sólo hay analogías muy remotas en el mundo animal. Y, con esto, el hombre ha creado un mundo nuevo de civilización, de saber, de crecimiento no-genético: de crecimiento

[4] [Para la explicación de Popper de las funciones del lenguaje, véanse *El desarrollo del conocimiento científico* [*C. & R.*], capítulos 4, 12; *Conocimiento objetivo* [*Obj. Kn.*], capítulos 2, 3, 4, 6; *Búsqueda sin término* [*U. Q.*], sección 15, y *El yo y su cerebro* [*S. & I. B.*], sección 17, con referencia especial a su discusión y ampliación de la explicación de Bühler. *Ed.*]

que no se transmite por el código genético; de crecimiento que depende no tanto de la selección natural como de la selección basada en la crítica racional.

Es, por tanto, al papel del lenguaje humano y del «Mundo 3» a donde tenemos que mirar cuando tratamos de explicar este tercer gran milagro: la emergencia del cerebro humano y de la mente humana, de la razón humana y de la libertad humana.

Determinismo e indeterminismo en física

El título de este ensayo es «El indeterminismo no basta»; es decir, no basta para la libertad humana. Pero tengo todavía que hacer, aunque sólo sea, un esbozo del determinismo clásico (determinismo físico, o determinismo del «Mundo 1») y del indeterminismo que se le opone. Además, todavía tengo que mostrar por qué esas dos ideas son insuficientes para una discusión de la libertad humana.

El determinismo clásico, o el determinismo del «Mundo 1», es una idea muy antigua, que fue formulada de la forma más precisa por Laplace sobre la base de la mecánica de Newton. [Véase la sección 10, más arriba.]

La tesis del determinismo de Laplace puede enunciarse de la manera siguiente: supongamos que nos dan las masas posiciones y velocidades exactas de todas las partículas materiales del universo en un momento del tiempo; entonces podremos, en principio, calcular con ayuda de la mecánica de Newton todo lo que ha ocurrido en el pasado y todo lo que ocurrirá en el futuro. Esto incluiría todos los movimientos físicos de todos los hombres y, por tanto, todas las palabras habladas o escritas, toda la poesía y toda la música que se escribirá jamás. El cálculo puede ser realizado por una máquina. *Sólo* necesita ser programada con las leyes del movimiento de Newton y con las condiciones iniciales existentes. Puede ser sorda como una tapia y totalmente desconocedora de los problemas de la composición musical. Pero será capaz de predecir los signos negros que colocará sobre papel pautado cualquier compositor dado, en el pasado o en el futuro.

Yo, personalmente, encuentro que el determinismo laplaciano no es una concepción convincente ni atractiva; y es un argumento dudoso, porque la calculadora tendría que exceder en mucho la complejidad del universo, como señaló (primero, creo) F. A.

Hayek [5]. Pero quizá merezca la pena resaltar que Laplace saca las conclusiones correctas de su idea de un «Mundo 1» causalmente cerrado y determinista. Si aceptamos la concepción de Laplace, entonces no tenemos que argüir (como lo hacen muchos filósofos) que estamos, no obstante, dotados de libertad y creatividad humanas auténticas.

Sin embargo, el determinismo laplaciano tuvo que ser modificado a raíz del colapso de unas tentativas que hizo Maxwell de reducir la electricidad y el magnetismo a la mecánica de Newton por medio de un modelo mecánico del éter. Con esas tentativas se derrumbó también la tesis de la clausura del «Mundo 1» mecánico de Newton: se hizo abierto hacia la parte electromagnética del «Mundo 1». A pesar de todo, Einstein, por ejemplo, continuó siendo determinista. Creyó, casi hasta el final de su vida, que una teoría determinista, unificada y cerrada, era posible, una teoría que comprendiese la mecánica, la gravitación y la electricidad. De hecho, la mayoría de los físicos se inclinan a considerar que un universo físico causalmente abierto (y, por tanto, indeterminista) —digamos que un universo físico que esté abierto a la influencia del «Mundo 2»— es una superstición típica, que sólo sostienen unos cuantos miembros de la Sociedad para la Investigación Física. Pocos físicos de renombre la tomarían en serio.

Pero otra forma de indeterminismo pasó a formar parte del credo oficial de los físicos. El nuevo indeterminismo que, introducido por la mecánica cuántica, supone la posibilidad de sucesos elementales fortuitos que son causalmente irreducibles.

Existen, al parecer, dos tipos de sucesos fortuitos. Uno de esos tipos se debe a la independencia de dos cadenas causales que por accidente interfieren en cierto lugar y tiempo y se combinan así para producir el suceso fortuito. Un ejemplo típico consiste en dos cadenas causales, una de las cuales hace que se suelte un ladrillo, mientras que la otra cadena causal independiente hace que un hombre se encuentre en tal posición que sea golpeado por el ladrillo. Este tipo de suceso fortuito (cuya teoría fue desarrollada por el propio Laplace, en su obra sobre la probabilidad) es perfectamente compatible con el determinismo laplaciano: cualquiera que estuviese de antemano en posesión de una información suficientemente amplia sobre los sucesos pertinentes podría haber

[5] F. A. von Hayek, *The Sensory Order*, 1952, capítulo 8, sección 6.

predicho lo que iba a ocurrir. Fue sólo *la incompletud de nuestro conocimiento* lo que dio lugar a este tipo de casualidad.

La mecánica cuántica, sin embargo, introdujo sucesos fortuitos de un segundo tipo, mucho más radical: casualidad absoluta. Según la mecánica cuántica, hay procesos físicos elementales que ya no son analizables en términos de cadenas causales, sino que consisten en los llamados «saltos cuánticos»; y se supone que un salto cuántico es un suceso absolutamente impredictible que no está controlado por las leyes causales ni por la coincidencia de leyes causales, sino solamente por las leyes de la probabilidad [6]. Así, la mecánica cuántica introdujo, a pesar de las protestas de Einstein, lo que él describió como «Dios-jugando-a-los-dados». La mecánica cuántica considera esos sucesos absolutamente casuales como los sucesos básicos del «Mundo 1». Los diversos resultados, tales como la desintegración de un átomo con la consiguiente emisión radiactiva, no están predeterminados y, por tanto, no pueden ser predichos por muy grande que sea nuestro conocimiento de todas las condiciones pertinentes previas al suceso. Pero podemos hacer predicciones estadísticas contrastables sobre esos procesos.

Aunque no creo que la mecánica cuántica permanezca como la última palabra en física, sí creo que su indeterminismo es fundamentalmente correcto. Yo creo que la mecánica clásica newtoniana es indeterminista en principio. Esto resulta claro si introducimos en ella modelos físicos de conocimiento humano; por ejemplo, computadoras [7]. La introducción de conocimiento humano objetivo en nuestro universo —la introducción del «Mundo 3» (no debemos olvidar que las computadoras, aunque no son humanas, están hechas por el hombre)— nos permite probar no sólo el carácter indeterminista de este universo, sino también su esencial apertura o incompletud.

Volviendo ahora a la mecánica cuántica, quiero señalar que el

---

[6] [Puede parecer que Popper concede aquí que hay saltos cuánticos mientras que en otros lugares parece estar de acuerdo con Schrödinger en que no los hay (véase *Teoría cuántica y el cisma en física* [*Q. T. & S. Ph.*], vol. III de este *Post Scriptum*, sección 13). Cuando se le preguntó sobre ello, Popper dijo que, aunque mantenía la imposibilidad de predecir los saltos excepto con criterios probabilísticos, estaba de acuerdo con Schrödinger en que es un problema abierto si debemos adoptar una interpretación del formalismo que considera instantáneos los saltos cuánticos. *Ed.*]

[7] [Véanse las exposiciones de Popper en *La miseria del historicismo* [*The Poverty*], Prefacio; «Indeterminism in Quantum Physics and in Classical Physics», en *British Journal for the Philosophy of Science I*, número 2, pp. 117-133, y número 3, páginas 173-195, y las secciones 20-22 de este tomo del *Post Scriptum*. *Ed.*]

indeterminismo de un Dios que juega a los dados, o de las leyes de la probabilidad, no crea espacio para la libertad humana. Porque lo que queremos entender no es sólo cómo podemos actuar *impredictiblemente y de una forma fortuita,* sino cómo podemos actuar *deliberada y racionalmente.* La famosa constancia probabilista de sucesos fortuitos como el de que se echen cartas al correo sin dirección puede ser una curiosidad interesante, pero no tiene semejanza alguna con el problema de la libertad de escribir un poema, bueno o malo, o de presentar una nueva hipótesis sobre, digamos, el origen del código genético.

Debe admitirse que si la mecánica cuántica está en lo cierto, entonces el determinismo laplaciano es erróneo y que los argumentos desde la física no pueden seguir usándose para combatir la doctrina del indeterminismo. Pero el indeterminismo no basta.

El indeterminismo no basta

Supongamos que el mundo está *parcial* pero no completamente determinado. Es decir, supongamos que los acontecimientos se suceden unos a otros según las leyes físicas, pero que hay a veces una cierta *holgura* en su conexión, que se rellena con secuencias impredictibles y quizá probabilistas similares a las que conocemos por la ruleta, los dados, por echar monedas a cara o cruz, o por la mecánica cuántica. Así, tendríamos un «Mundo 1» indeterminista, como, efectivamente, ya he sugerido. *Pero nosotros no ganamos nada si este «Mundo 1» está causalmente cerrado en relación con el «Mundo 2» y el «Mundo 3».* Un «Mundo 1» indeterminista de este modo sería impredictible; sin embargo, el «Mundo 2» y, con él, el «Mundo 3» no podrían tener ninguna influencia sobre él. Un «Mundo 1» indeterminista *cerrado* continuaría su curso como antes, fueran los que fuesen nuestros sentimientos o nuestros deseos, con la única diferencia respecto al mundo de Laplace de que no podríamos predecirlo, incluso aunque conociésemos *todo* lo referente a su estado actual: sería un mundo dominado, aunque fuese sólo parcialmente, por la suerte.

Así pues, el indeterminismo es *necesario, pero insuficiente,* para permitir la libertad humana y, en especial, la creatividad. Lo que realmente necesitamos es la tesis de que *el «Mundo 1» es incompleto;* que puede ser influido por el «Mundo 2»; que puede

interactuar con el «Mundo 2», o que está causalmente *abierto* hacia el «Mundo 2», y de ahí también hacia el «Mundo 3».

Volvemos así a nuestro argumento central: tenemos que postular que el «Mundo 1» no sea autónomo o «cerrado», sino abierto hacia el «Mundo 2»; que pueda ser influido por el «Mundo 2» de la misma manera que el «Mundo 2» puede ser influido por el «Mundo 3» y, naturalmente, por el «Mundo 1».

Determinismo y naturalismo

Poca duda puede haber de que el motivo filosófico fundamental en favor del determinismo laplaciano y de que la teoría de que el «Mundo 1» está cerrado causalmente estriba en la comprensión de que el hombre es un animal y en el deseo de vernos a nosotros mismos como parte de la naturaleza. Creo que el motivo es correcto: si la naturaleza fuese totalmente determinista, entonces también lo sería el reino de las acciones humanas; de hecho no habría acciones, sino, como mucho, la apariencia de acciones.

Pero puede darse la vuelta al argumento. Si el hombre es libre, al menos en parte, también lo es la naturaleza; y el «Mundo 1» físico es abierto. Y hay muchas razones para considerar al hombre al menos parcialmente libre. La concepción contraria —la de Laplace— lleva a la predestinación. Lleva a la percepción de que hace miles de millones de años las partículas elementales del «Mundo 1» contenían la poesía de Homero, la filosofía de Platón y las sinfonías de Beethoven de la misma manera que una semilla contiene una planta; que la historia está predestinada y, con ella, todos los actos de la creatividad humana. Y la versión de esa idea que proporciona la teoría cuántica es igualmente mala. Si tiene alguna relación con la creatividad humana, entonces hace de ella una cuestión de pura suerte. Sin duda hay en ella un elemento de suerte. Pero la teoría de que la creación de obras de arte o música puede explicarse, en última instancia, en términos de química o física, me parece absurda. En la medida que puede explicarse la creación de música, tiene que hacerse, al menos en parte, en términos de la influencia de otra música (que también estimula la creatividad en el músico); y, lo más importante, en términos de la estructura interna, las leyes y las restricciones internas, que juegan un papel en la música y en todos los demás fenómenos del «Mun-

do 3»; leyes y restricciones cuya absorción (y cuyo ocasional desafío) es enormemente importante para la creatividad del músico.

Así, nuestra libertad, y especialmente nuestra libertad de crear, está claramente limitada por las restricciones de los tres mundos. Si Beethoven, por alguna desgracia, hubiera sido sordo de nacimiento, difícilmente hubiera podido llegar a ser compositor. Como compositor, subordinó libremente su libertad a las restricciones estructurales del «Mundo 3». El «Mundo 3» autónomo fue el mundo en el que él hizo sus auténticos grandes descubrimientos, siendo libre de escoger su camino como un descubridor en el Himalaya, pero restringido por el camino escogido hasta el momento y por las restricciones del mundo que iba descubriendo. (Se podría decir lo mismo sobre Gödel.)

El universo abierto

Esto nos vuelve a llevar a la aserción de que hay interacción entre los «Mundos 1, 2 y 3».

No tengo la menor duda de que los «Mundos 2 y 3» interactúan. Si tratamos de captar o entender una teoría, o de recordar una sinfonía, entonces nuestras mentes están influidas causalmente; no simplemente por el recuerdo de los sonidos almacenados en nuestro cerebro, sino, al menos parcialmente, por la obra del compositor, por las estructuras internas autónomas del objeto del «Mundo 3» que tratamos de captar.

Todo esto significa que el «Mundo 3» puede actuar sobre el «Mundo 2» de nuestras mentes. Pero, si ocurre así, no hay duda de que, cuando un matemático *escribe* sus resultados, del «Mundo 3» en papel (físico), su mente —su «Mundo 2»— actúa sobre el «Mundo 1». Así, el «Mundo 1» está abierto sobre el «Mundo 2», de la misma manera que el «Mundo 2» está abierto hacia el «Mundo 3».

Esto tiene una importancia fundamental, porque muestra que la naturaleza, o el universo al que pertenecemos, y que contiene como partes los «Mundos 1, 2 y 3», es, él mismo, abierto; contiene el «Mundo 3», y puede mostrarse que el «Mundo 3» es *intrínsecamente abierto*.

Un aspecto de la apertura del «Mundo 3» es una consecuencia del teorema de Gödel según el cual la aritmética axiomatizada no es completable. Pero la incompletabilidad y la apertura del uni-

verso se comprenden mejor con una versión de la conocida historia del hombre que dibuja un mapa de su habitación incluyendo en el mapa el mapa que está dibujando. Su tarea no puede ser completada, porque tiene que dibujar en su mapa su última anotación.

El cuento del mapa es un caso trivial comparado con las teorías del «Mundo 3» y su impacto sobre el «Mundo 1», aunque ilustra, de manera sencilla, la incompletud de un universo que contiene los objetos del conocimiento del «Mundo 3». Pero, por el momento, no ilustra el indeterminismo. Porque cada uno de los distintos «últimos» trazos realmente dibujados en el mapa determina, dentro de la infinita sucesión de anotaciones que hay que hacer, una anotación determinada. Sin embargo, esta determinación de los trazos es válida sólo si no consideramos la falibilidad de todo conocimiento humano (una falibilidad que juega un papel considerable en los problemas, teorías y errores del «Mundo 3»). Tomando esto en cuenta, cada uno de esos «últimos» trazos anotados en el mapa constituye un nuevo *problema* para el dibujante, el problema de dibujar un trazo *más* que describa el último trazo *con precisión*. A causa de la falibilidad que caracteriza todo conocimiento humano, este problema no puede ser resuelto por el dibujante con precisión absoluta; y cuanto más pequeños son los trazos que va haciendo el dibujante, mayor será la imprecisión relativa, que será, en principio, impredictible e indeterminada y que aumentará constantemente. De esta forma, el cuento del mapa muestra cómo la falibilidad que afecta al conocimiento humano objetivo contribuye al esencial indeterminismo y a la apertura de un universo que contiene, como parte de sí mismo, el conocimiento humano.

Así, el universo tiene que ser abierto si contiene conocimiento humano; artículos y libros, como éste, que son, por un lado, objetos del «Mundo 1» y, por el otro, objetos del «Mundo 3» que tratan faliblemente de enunciar o describir el falible conocimiento humano.

Así pues, vivimos en un universo abierto. No podíamos hacer este descubrimiento antes de que hubiera conocimiento humano. Pero, una vez que lo hemos hecho, no hay razón para pensar que la apertura depende exclusivamente de la existencia de conocimiento humano. Es mucho más razonable rechazar todas las teorías de un universo cerrado: la de un universo cerrado causalmente tanto como la de uno cerrado probabilísticamente, recha-

zando, por tanto, el universo cerrado concebido por Laplace y el contemplado por la mecánica de ondas. Nuestro universo es en parte causal, en parte probabilista y en parte abierto: es emergente. La concepción contraria proviene de confundir el carácter de nuestras teorías del «Mundo 3», hechas por el hombre, sobre el «Mundo 1» —especialmente sus características ultra-simplificaciones—, con el propio carácter del «Mundo 1». Confusión en la que no se debería incurrir.

Hasta el momento, no se ha ofrecido ninguna buena razón en contra de la apertura de nuestro universo o contra el hecho de que constantemente están emergiendo de él cosas radicalmente nuevas; y hasta el momento no se han ofrecido buenas razones para poner en duda la libertad y la creatividad humanas, una creatividad que está restringida a la vez que inspirada por la estructura interna del «Mundo 3».

Ciertamente, el hombre forma parte de la naturaleza, pero al crear el «Mundo 3» se ha transcendido a sí mismo y a la naturaleza tal como existía antes de él. Y la libertad humana es, desde luego, parte de la naturaleza, mas transciende la naturaleza, al menos tal como ésta existía antes de la emergencia del lenguaje humano y del pensamiento crítico y del saber humano.

El indeterminismo no basta: para entender la libertad humana necesitamos más; necesitamos la apertura del «Mundo 1» hacia el «Mundo 2»; y del «Mundo 2» hacia el «Mundo 3», y la autónoma e intrínseca apertura del «Mundo 3», el mundo de los productos de la mente humana y, especialmente, del saber humano.

REDUCCION CIENTIFICA Y LA INCOMPLETUD ESENCIAL DE TODA CIENCIA [1]

I

El determinismo está, históricamente, estrechamente relacionado con la tesis del «reduccionismo». Un determinista «científico», en el sentido presentado en este libro, tiene que ser reduccionista, aunque no es necesario que todo reduccionista sea determinista. En este *Addendum* quiero hablar brevemente del reduccionismo.

Las cuestiones más destacadas del reduccionismo son tres, creo yo:

1) ¿Podemos reducir, o tener la esperanza de llegar a reducir, la biología a física o a física y química?

2) ¿Podemos reducir a biología, o tener la esperanza de llegar a reducir a biología, aquellas experiencias conscientes subjetivas que podemos atribuir a los animales y, si la pregunta 1) se contesta afirmativamente, podemos reducirlas, aún más, a física y química?

3) ¿Podemos reducir, o tener la esperanza de llegar a reducir, la conciencia del yo y la creatividad de la mente humana a una experiencia animal, y así, si las preguntas 1) y 2) se contestan afirmativamente, a física y química?

Es obvio que las respuestas a esas tres preguntas (sobre las que volveré más adelante) dependerán en parte del significado de la palabra «reducir». Pero, por las razones que ya he dado en otro lugar, me opongo al método de analizar significados y al intento de resolver problemas serios por medio de definiciones. En cambio, lo que me propongo hacer es esto:

Empezaré por exponer algunos ejemplos de reducciones que han tenido éxito y de otras que no lo han tenido, en las diversas

---

[1] [Este *Addendum* no formaba parte del *Post Scriptum* original, y es una versión revisada de un ensayo publicado originariamente en F. J. Ayala y T. Dobzhansky, eds.: *Studies in the Philosophy of Biology*, 1974. *Ed.*]

ciencias, y hablaré especialmente de la reducción de la química a física; y también de los residuos que han dejado esos programas de investigación reduccionistas.

En el transcurso de esta exposición defenderé tres tesis. En primer lugar, sugeriré que los científicos tienen que ser reduccionistas en el sentido de que nada obtiene un éxito tan grande en la ciencia como una reducción que lo haya alcanzado (tal como la reducción de Newton a —o más bien su explicación promedio de [2]— su teoría de las leyes de Kepler y de Galileo, y la corrección que hizo de ellas). Una reducción de éxito es, quizá, la forma de mayor éxito que puede concebirse entre todas las explicaciones científicas, puesto que logra lo que subrayó Meyerson: una *identificación* de lo desconocido con lo conocido [3]. Sin embargo, al contrario que una reducción, una explicación, con ayuda de una nueva teoría, explica lo conocido —el problema conocido— por medio de algo desconocido: por medio de una nueva conjetura [4].

En segundo lugar, sugeriré que los científicos *tienen* que aceptar de buen grado el reduccionismo como *método:* o bien tienen que ser unos ingenuos o, si no, tienen que ser reduccionistas más o menos críticos; de hecho, unos reduccionistas críticos un poco a la desesperada, argüiré, porque apenas ha habido alguna reducción científica de importancia que haya tenido un éxito *total:* casi siempre hay un residuo no resuelto incluso en los programas de investigación reduccionista de mayor éxito.

En tercer lugar, sostendré que no parece haber ningún buen argumento en favor del reduccionismo *filosófico,* mientras que, por el contrario, hay muy buenos argumentos contra el esencialismo con el que el reduccionismo filosófico parece estar estrechamente ligado [5]. Deberíamos, no obstante, continuar intentando

---

[2] [Véanse *Conocimiento objetivo (Obj. Kn.),* capítulo 5, y *Realismo y el objetivo de la ciencia (R. & A. of Sc.),* tomo I del *Post Scriptum,* sección 15. *Ed.*]

[3] E. Mayerson, *Identity and Reality,* 1980. Traducción al inglés de 1930.

[4] Véase *El desarrollo del conocimiento científico [C. & R.*].

[5] Aquí, paso por alto —quizá por descuido o porque me disgustan las minucias terminológicas— la distinción que muy bien podría hacerse entre explicaciones en general y «reducción» en el sentido de explicación por medio de una teoría mejor contrastada o más «fundamental». Una distinción de mayor interés, supongo, sería la que hay entre una explicación de algo conocido por una teoría nueva (desconocida), por un lado, y una reducción a una teoría antigua (conocida), por otro. Supongo que podría hacerse una distinción entre una *reducción* que explica una teoría por una teoría existente y una *explicación con ayuda de una nueva teoría:* aunque no voy a pelear por cuestiones de palabras, no me sentiría inclinado a llamar «reducción» a una explicación que se sirve de una nueva teoría. Sin embargo,

en el terreno metodológico las reducciones. Porque podemos aprender muchísimo incluso de los intentos de reducción incompletos o sin éxito, y los problemas que quedan abiertos de esta manera constituyen nuestras posesiones intelectuales más valiosas: un mayor énfasis sobre lo que a menudo se consideran nuestros fracasos científicos (o, dicho de otro modo, sobre los grandes problemas abiertos de la ciencia) puede hacernos mucho bien.

II

Aparte de la de Newton, una de las pocas reducciones que yo conozco que hayan tenido éxito casi completo es la reducción de las fracciones racionales a pares ordenados de números naturales (es decir, a relaciones o proporciones entre ellos). Aunque lo hicieron los griegos, puede decirse que incluso esa reducción dejó un *residuo* que no se abordó hasta el siglo XX, con la reducción, de éxito, de Wiener (1914) y Kuratowski (1920), del par ordenado a un par no ordenado de pares no ordenados; además, uno debería percatarse de que la reducción es una reducción a *conjuntos* de pares equivalentes, más bien que a pares sin más. Esta reducción estimuló el programa cosmológico de investigación de aritmetización emprendido por la escuela pitagórica [6], que, sin embargo, se desbarató con la prueba de la exigencia de irracionales tales como las raíces cuadradas de 2, 3 ó 5 [7]. Platón sustituyó el programa cosmológico de investigación de aritmetización por un programa de geometrización y este programa se llevó adelante con éxito

---

si se adopta esta terminología, la explicación de la teoría ondulatoria de la propagación de la luz, por medio de la teoría de Maxwell del electromagnetismo, podría considerarse como un ejemplo de una reducción de éxito completo (quizá el único ejemplo de una reducción de total éxito). Sin embargo, puede ser mejor no describirla como una reducción de una teoría a otra, o de una parte de la física a otra, sino, más bien, como una teoría radicalmente nueva que consiguió unificar dos partes de la física.

[Para la discusión que hace Popper del esencialismo en este contexto, y para la conexión de éste con el justificacionismo, véase *Realismo y el objetivo de la ciencia (R. & A. of Sc.)*, tomo I del *Post Scriptum*, sección 2; *Búsqueda sin término (U. Q.)*, sección 7, y *La sociedad abierta y sus enemigos (O. S.)*, capítulo 11. *Ed.*]

[6] [Para la idea de Popper sobre los programas de investigación, véase *Teoría cuántica y el cisma en Física (Q. T. & S. Ph.)*, vol. III del *Post Scriptum*. Epílogo metafísico. *Ed.*]

[7] Véanse *La sociedad abierta y sus enemigos (O. S.)*, vol. I, capítulo 6, número 9, y *El desarrollo del conocimiento científico (C. & R.)*, capítulo 2, pp. 80-115.

desde Euclides hasta Einstein. Sin embargo, la invención del cálculo por Newton y Leibniz (y del problema de la exclusión de algunos resultados paradójicos que sus propios métodos intuitivos no excluyeron) creó la necesidad de una nueva aritmetización, una nueva reducción a los números naturales. A pesar de los espectaculares éxitos del siglo XIX y los primeros años del XX, esta reducción no ha tenido un éxito completo.

Por mencionar solamente uno de los residuos no resueltos, una reducción a una sucesión de números naturales o a un conjunto en el sentido de la moderna teoría de conjuntos no es lo mismo, ni siquiera parecido, a una reducción, a un conjunto de pares ordenados equivalentes de números naturales. Mientras la idea de conjunto se utilizó de forma ingenua y puramente intuitiva (como hizo Cantor), esto no fue quizá obvio. Pero las paradojas de conjuntos infinitos (tratadas por Bolzano, Cantor y Russell) y la necesidad de axiomatizar la teoría de conjuntos mostraron, por decir lo mínimo, que la reducción lograda no era una aritmetización franca —una reducción a números naturales—, sino una reducción a la teoría de conjuntos axiomática; y ésta resultó ser una empresa sumamente sofisticada y algo peligrosa.

Para resumir este ejemplo, el programa de la aritmetización —es decir, la reducción de la geometría y de los irracionales a números naturales— ha fracasado en parte. Pero el número de problemas inesperados y la cantidad de conocimiento inesperado a que ha dado lugar este fracaso son abrumadores. Esto puede generalizarse: incluso cuando no tenemos éxitos como reduccionistas, el número de resultados interesantes e inesperados que podemos obtener en el camino hacia nuestro fracaso pueden ser de enorme valor.

III

He mencionado el fracaso parcial del intento de reducción de los irracionales a números naturales y he indicado también que los programas de reducción forman parte de las actividades de la explicación, simplificación y comprensión científicas y matemáticas.

Trataré ahora un poco más detalladamente algunos éxitos y algunos fracasos de programas de reducción en física, y, en particular, los éxitos parciales desde la reducción de la macrofísica

a microfísica y de la química tanto a macrofísica como a microfísica.

Uso el término «explicación última» para el intento de explicar o reducir las cosas apelando a algo que no necesita ser explicado o que no puede serlo, especialmente una «esencia» o «sustancia» *(ousia)* [8].

Un ejemplo llamativo es la reducción cartesiana de la totalidad de la física de los cuerpos inanimados a *sustancia extensa;* una sustancia (materia) con una sola propiedad esencial, a saber, la extensión espacial [9].

Esta tentativa de reducir la totalidad de la física a una propiedad aparentemente esencial de la materia tuvo mucho éxito en la medida en que hizo surgir una visión inteligible del universo físico. El universo físico cartesiano era un mecanismo de relojería en movimiento, de vértices en el que cada «cuerpo» o «parte de materia» empujaba a la parte vecina y era empujado por su vecina del lado contrario. Sólo materia podía encontrarse en el mundo físico y todo el espacio estaba lleno de ella. De hecho, también el espacio se reducía a materia, ya que no había espacio vacío, sino sólo la extensión espacial esencial de la materia. Y sólo había un modo, puramente físico, de causación: *toda causación era empuje* o acción por contacto.

Esta forma de mirar el mundo era considerada satisfactoria, incluso por Newton, aunque se sintió impelido a introducir, con su teoría de la gravitación, un nuevo tipo de causación: *la atracción,* o acción a distancia.

El casi increíble éxito explicativo y predictivo de la teoría de Newton destruyó el programa de reducción cartesiana. El propio Newton intentó realizar el programa cartesiano de reducción explicando la atracción gravitatoria por el «impulso» (presión de radiación combinado con el efecto sombrilla) del bombardeo de una partícula cósmica (el intento va normalmente unido al nombre de LeSage) [10]. Pero yo creo que Newton se dio cuenta de la objeción fatal contra esta teoría. Cabe suponer que reduciría la atracción y la acción a distancia a impulso, y acción por contacto; pero eso significaría también que todos los cuerpos móviles se

---

[8] Véase *El desarrollo del conocimiento científico (C. & R.),* capítulo 3, páginas 116-141.

[9] Véanse *El yo y su cerebro* [*S. & I. B.*], capítulos P1 y P3, y *Teoría cuántica y el cisma en Física* [*Q. T. & S. Ph.*], vol. III del *Post Scriptum.* Epílogo metafísico.

[10] Véase *El desarrollo del conocimiento científico* [*C. & R.*].

moverían en un medio resistente que actuaría como freno de su movimiento (considérese el exceso de empuje de la lluvia en el parabrisas de un coche sobre el empuje que recibe la ventana posterior) y que invalidaría así el uso que Newton hacía de la ley de inercia. Así, a pesar de su atractivo intuitivo y a pesar del rechazo como «absurda» por parte del propio Newton de la idea de que la atracción a distancia podía ser una propiedad esencial de la materia, la tentativa de una reducción última de la atracción a empuje se desbarata.

Este es un sencillo ejemplo de una reducción física prometedora y de su fracaso, y de lo mucho que puede aprenderse al intentar una reducción y al descubrir que, y quizá incluso por qué, falla.

(Conjeturo que este fracaso fue la razón inmediata por la que Newton describió el espacio como el sensorio de Dios. El espacio «se percataba», por decirlo así, de la distribución de todos los cuerpos: era, en un sentido, omnisciente. Era también omnipresente, porque transmitía su conocimiento con infinita velocidad a todos los lugares en cada momento. Así, el espacio, compartiendo como mínimo dos propiedades características de la esencia divina, era él mismo parte de la esencia divina. Este, sugiero yo, era otro intento de Newton de dar una explicación última esencialista.)

La reducción cartesiana ilustra por qué, por razones metodológicas, debemos intentar las reducciones. Pero puede indicar también por qué no debemos ser confiados, sino que debemos, en cierto modo, desesperar del éxito completo de nuestros intentos de reducción.

IV

El intento cartesiano de reducir todo lo del mundo físico a extensión y empuje fue un fracaso cuando se compara con el éxito de la teoría de la gravitación de Newton. Ese éxito fue tan grande que los newtonianos, empezando por Roger Cotes, empezaron a considerar la teoría de Newton como una explicación última y, así, consideraron la *atracción gravitatoria* como una propiedad esencial de la materia a pesar de las opiniones contrarias del propio Newton. Pero Newton no había visto ninguna razón por la que la *extensión* (de sus átomos) y la *inercia* no pudieran ser pro-

piedades esenciales de la masa [11]. Así pues, Newton era claramente consciente de la distinción, en la que más tarde hizo hincapié Einstein, entre masa inerte y masa gravitatoria, y del problema que ponía de manifiesto su proporcionalidad (o igualdad); un problema que, a causa del oscurantismo del enfoque esencialista, casi se perdió de vista entre Newton y Eötvös o, incluso, Einstein.

La teoría especial de la relatividad de Einstein destruyó la identidad esencialista entre masa inerte y masa gravitatoria, y por eso intentó él explicarla, algo *ad hoc,* por medio de su principio de equivalencia. Pero cuando se descubrió (el primero en hacerlo fue Cornelius Lanczos) que las ecuaciones de la gravitación de Einstein llevaban por sí solas al principio, que anteriormente se había expuesto separadamente, de que los cuerpos gravitatorios se mueven en una geodésica de espacio-tiempo, el principio de inercia se redujo, de hecho, a las ecuaciones de la gravitación y, de esta manera, la masa inerte a masa gravitatoria. (Yo creo que Einstein, aunque fuertemente impresionado por la importancia de este resultado, no aceptó plenamente que con él se resolviera el problema central de Mach —la explicación de la inercia— de una forma más satisfactoria que el famoso pero ambiguo «principio de Mach»: el principio de que la inercia de cada cuerpo se debe al efecto combinado de todos los cuerpos restantes del universo. Para desilusión de Einstein, ese principio era, al menos según algunas de las interpretaciones que se le daban, incomparable con la relatividad general que, para un espacio vacío de todos los cuerpos, da lugar a la relatividad especial, en la que la ley de inercia, contrariamente a la sugerencia de Mach, sigue siendo válida.)

Este es un ejemplo sumamente satisfactorio de una reducción con éxito: la reducción de un principio de inercia generalizado a un principio de gravitación generalizado. Pero rara vez se ha considerado desde esa óptica; ni siquiera lo hizo Einstein, aunque sintió con fuerza la importancia de un resultado que, desde un punto de vista puramente matemático, podía considerarse elegante, pero no particularmente importante. Porque la dependencia o independencia de un axioma dentro de un sistema de axiomas no tiene, en general, más que un interés formal. ¿Por qué iba a importar, por tanto, que la ley del movimiento en una geodésica tuviera que ser tomada como un axioma distinto o que pudiera derivarse del resto de la teoría de la gravedad? La respuesta es que, por me-

[11] Véase *El desarrollo del conocimiento científico* [*C. & R.*].

dio de su derivación, la identidad de la masa inerte y la masa gravitatoria quedaba *explicada* y la primera reducida a la segunda.

De este modo, puede decirse que el gran problema de Newton de la acción a distancia (explicado con la fraseología del esencialismo) no fue resuelto tanto por la velocidad finita de la acción gravitatoria de Einstein como por la reducción de la materia inerte a materia gravitatoria.

V

Newton y los newtonianos sabían, naturalmente, de la existencia de fuerzas magnéticas y eléctricas; y, hasta por lo menos el principio del siglo XX, se hicieron tentativas de reducir la teoría electromagnética a la mecánica de Newton o a una forma modificada de ella.

El problema más destacado de esta tentativa era la reducción de fuerzas no-centrales *prima facie* (fuerzas de Oersted) a fuerzas centrales, las únicas que parecían encajar en una teoría newtoniana, incluso modificada. Los nombres sobresalientes en estos acontecimientos fueron Ampère y Weber.

También Maxwell empezó a intentar reducir el campo electromagnético de Faraday de (líneas de) fuerzas a un mecanismo newtoniano o modelo del éter luminífero. Pero renunció al intento (aunque no al éter luminífero como portador del campo electromagnético). Helmholtz también se sintió atraído por un programa de reducción newtoniano y parcialmente cartesiano, y cuando sugirió a su discípulo Heinrich Hertz que debería trabajar sobre ese problema, Helmholtz parece haberlo hecho con la esperanza de salvar el programa de investigación de mecánica. Pero aceptó la confirmación de Hertz de que las ecuaciones de Maxwell eran cruciales. Después de Hertz y de J. J. Thomson, se hizo más atractivo precisamente el programa de investigación opuesto: el programa de reducir la mecánica a teoría electromagnética.

La teoría electromagnética de la materia —es decir, la reducción tanto de la mecánica como de la química a una teoría electromagnética del atomismo— tuvo un sorprendente éxito desde, por lo menos, 1912, el año del modelo planetario de Rutherford o del átomo nuclear, hasta 1932, aproximadamente.

De hecho, la mecánica cuántica (o «la nueva teoría cuántica», como se la llamó durante cierto tiempo) fue, por lo menos hasta

1935, simplemente otro nombre para lo que entonces se consideraba la forma final de la reducción de la mecánica a la nueva *teoría electromagnética de la materia.*

Para comprender lo importante que les pareció esta teoría a los principales físicos, incluso poco antes de la mecánica cuántica, voy a citar a Einstein, que escribió [12]:

> ... según nuestras concepciones actuales, las partículas elementales [es decir, electrones y protones] no son... *nada más* que condensaciones del campo electromagnético..., nuestra... visión del universo presenta dos realidades..., a saber, éter gravitatorio y campo electromagnético, o —como podía llamarse también— espacio y materia.

El «nada más» que he subrayado es característico de la reducción a lo grande. De hecho, hasta el final de su vida, Einstein trató de unificar los campos gravitatorios y electromagnéticos en una teoría de campo unificada, incluso después de que sus opiniones de 1920 hubieran sido sustituidas, o, más bien, invalidadas (debido, especialmente, al descubrimiento de fuerzas nucleares).

Lo que equivale, esencialmente, a la misma concepción reduccionista era aceptado en aquel momento (1932) por casi todos los físicos principales: Eddington y Dirac en Inglaterra, y, además de Einstein, Bohr, De Broglie, Schrödinger, Heinsenberg, Born y Pauli en el continente europeo. Y Robert A. Millikan, entonces en el Instituto de Tecnología de California, formuló un enunciado impresionante de esa concepción [13].

> Realmente, nada extraordinariamente simplificador había ocurrido nunca en la ciencia como toda la serie de descubrimientos que culminó alrededor de 1914, año en que por fin se aceptó prácticamente por todos la teoría de que el mundo material no contiene más que dos entidades fundamentales, a saber, electrones negativos y electrones positivos, exactamente iguales en carga, pero muy diferentes en masa: el electrón positivo —que ahora suele llamarse protón—, 1.850 veces más pesado que el negativo, que ahora suele llamarse simplemente electrón.

Este pasaje reduccionista se escribió en el momento preciso: fue en ese mismo año (1932) cuando Chadwick publicó su descu-

---

[12] A. Einstein, *Äther und Relativitätstheorie,* 1922. Véase también mi «Quantum Mechanics without 'the Observer'», en Mario Bunge, ed., *Quantum Theory and Reality,* 1967, pp. 7-44, reimpreso ahora, en versión revisada, como Introducción a *Teoría cuántica y el cisma en Física* [*Q. T. & S. Ph.*], vol. III del *Post Scriptum.*

[13] R. A. Millikan, *Time, Matter and Values,* 1932, p. 46.

brimiento del neutrón, y Anderson descubrió por primera vez el positrón. Aun así, algunos de los más grandes físicos, como Eddington (1936), seguían creyendo, incluso después de la sugerencia de Yukawa de la existencia de lo que iba a llamarse el mesón (1935), que con el advenimiento de la mecánica cuántica la teoría electromagnética de la materia había entrado en su estado final y que toda la materia consistía en electrones y protones [14].

En efecto, la reducción de la mecánica y la química a la teoría electromagnética de la materia parecía casi perfecta. Lo que Descartes y Newton concibieron como la esencia de la materia que llena el espacio y como empuje cartesiano había sido reducido (según lo había postulado Leibniz hacía ya mucho tiempo) a *fuerzas repulsivas*, las fuerzas ejercidas por los electrones negativos sobre los electrones negativos. La neutralidad eléctrica de la materia se explicaba por el número igual de protones positivos y de electrones negativos, y la electrificación (ionización) de la materia se explicaba por una pérdida de electrones (o por exceso de electrones) en la corteza eléctrica planetaria del átomo.

La química se había reducido a física (o al menos eso parecía) por medio de la teoría cuántica de Bohr del sistema periódico de los elementos, una teoría que fue ingeniosamente perfeccionada con el uso del principio de exclusión de Pauli; y la teoría de la composición química, y de la naturaleza de los enlaces covalentes químicos, fue reducida por Heitler y London (1927) a una teoría de valencia (homeopolar) que también hacía uso del principio de Pauli.

Aunque se reveló que la materia era una estructura compleja en vez de una sustancia irreducible, nunca antes había habido tanta unidad en el universo de la física, o tal grado de reducción.

Y nunca ha vuelto a lograrse desde entonces.

Cierto, seguimos creyendo en la reducción del empuje cartesiano a fuerzas electromagnéticas; y la teoría de Bohr del sistema periódico de los elementos, aunque muy cambiada por la introducción de isótopos, ha sobrevivido en gran medida. Pero todo lo demás en esta bella reducción del universo a un universo electromagnético con dos partículas como bloques estables de construcción se ha desintegrado ya. Indudablemente, hemos aprendido

[14] J. Chadwick, «Possible existence of a neutron», *Nature*, 129, p. 132; C. D. Anderson, «Cosmic Ray Bursts», *Physical Review*, 43, pp. 368-9, y «The positive electron», *Physical Review*, 43, pp. 491-4; A. Eddington, *Relativity Theory of Protons and Electrons*, 1936, y H. Yukawa, «On the interaction of elementary particles», *Proceedings of the Phisico-Mathematical Society of Japan*, 3.ª serie, 17, páginas 48-57.

una cantidad inmensa de hechos nuevos en el proceso de esta desintegración: ésta es una de mis tesis principales. *Pero la simplicidad, y el programa de reducción, han desaparecido.*

Este proceso, que se inició con el descubrimiento de los neutrones y los positrones, ha continuado con el descubrimiento, desde entonces, de nuevas partículas elementales. Pero no es siquiera la teoría de partículas la mayor dificultad. El verdadero destrozo se debe al descubrimiento de nuevos tipos de fuerzas, especialmente de fuerzas nucleares de corto alcance, aparentemente irreducibles a fuerzas electromagnéticas y a fuerzas gravitatorias.

Las fuerzas gravitatorias no preocupaban mucho a los físicos en aquellos momentos, porque la relatividad general había encontrado una explicación para prescindir de ellas y se tenía la esperanza de que el dualismo de las fuerzas gravitatorias y electromagnéticas sería reemplazado por una teoría de campo unificada. Pero tenemos por lo menos cuatro tipos muy diferentes, y todavía irreducibles, de fuerzas en física: fuerzas de gravitación, fuerzas de interacción débil de desintegración, fuerzas electromagnéticas y fuerzas nucleares [15].

VI

Así pues, la mecánica cartesiana —considerada en un tiempo por Descartes y Newton como la base a la que había que reducir todo lo demás— se reducía, y todavía se reduce con éxito, a electromagnetismo. Pero ¿qué hay de la que se ha considerado como la más impresionante reducción, la de la química a física cuántica?

Supongamos, en pro del argumento, que tenemos una reducción plenamente satisfactoria a la teoría cuántica de los enlaces

[15] Existen hoy nuevos y más amplios intentos de lograr una teoría de campo unificada, intentos de unificar las cuatro fuerzas naturales. Entre ellos, podemos mencionar la obra de Steven Weinberg y Abbus Salan, lo que a menudo se llama «la Teoría Weinberg-Salam de la Interacción Débil». Esta teoría propuso que las fuerzas electromagnéticas y las fuerzas nucleares débiles (que causan la desintegración radiactiva en algunos tipos de núcleos atómicos) son aspectos o facetas del mismo fenómeno; y ha habido alguna corroboración a esta propuesta por medio de contrastaciones de corriente neutra en núcleos atómicos. Pero las fuerzas gravitatorias y las fuerzas nucleares fuertes siguen estando fuera del alcance de esta teoría y la propia teoría es, por tanto, en ciertos aspectos, todavía un programa metafísico de investigación.

químicos (tanto de enlaces covalentes o de electrones gemelos como de enlaces no covalentes; por ejemplo, enlaces de enchufe-y-agujero) a pesar de la reveladora observación de Pauling (1959), autor de *The Nature of the Chemical Bond*, de que era incapaz de «definir» (o enunciar con precisión) cuál era la naturaleza del enlace químico. Supongamos, aún más, también en pro del argumento, que tenemos una teoría plenamente satisfactoria de las fuerzas nucleares, del sistema periódico de los elementos y sus isótopos y, especialmente, de la estabilidad e inestabilidad de los núcleos más pesados. ¿Constituye esto una reducción plenamente satisfactoria de la química a la mecánica cuántica?

Yo no lo creo. Tiene que introducirse una idea enteramente nueva, una idea que es de algún modo ajena a la teoría física: la idea de evolución, de la historia de nuestro universo, de la cosmogonía e, incluso más, de la cosmogenia.

La razón de ello es que la tabla periódica de los elementos y la teoría (reformulada) del sistema periódico de Bohr explican los núcleos más pesados, diciendo que están compuestos de otros más ligeros; y, en último término, que están compuestos de núcleos de hidrógeno (protones) y de neutrones (que a su vez *podrían* considerarse un tipo de composición de protones y electrones). Y esta teoría supone que los elementos más pesados tienen una historia, que las propiedades de sus núcleos resultan de hecho de un raro proceso que hace que varios núcleos de hidrógeno se fundan en núcleos más pesados, en condiciones que se encuentran sólo raramente en el cosmos.

Tenemos amplia evidencia en favor de la opinión de que esto ocurrió realmente y que sigue ocurriendo; que los elementos más pesados tienen una historia evolutiva exclusiva y que el proceso de fusión por el cual el hidrógeno pesado se transforma en helio es la principal fuente de energía de nuestro propio Sol y también de la bomba de hidrógeno. Así, el helio y todos los elementos más pesados son el resultado de la evolución cosmológica. Su historia, y en especial la historia de los elementos más pesados, es, según las concepciones cosmológicas actuales, una historia extraña. Los elementos más pesados se consideran actualmente productos de explosiones de supernovas. Como el helio, según ciertos cálculos recientes, forma el 25 por 100 de toda la materia por masa y el hidrógeno dos tercios o tres cuartos de toda la materia por masa, todos los núcleos más pesados parecen ser extremadamente raros (todos juntos, quizá 1 ó 2 por 100 por masa). Así, la Tierra y, pre-

sumiblemente, los demás planetas de nuestro sistema solar están hechos principalmente de materiales muy raros (y yo diría que preciosos).

En la actualidad, la teoría más ampliamente aceptada sobre el origen del universo [16] —la del tórrido *big bang*— sostiene que la mayor parte del helio es producto del propio *big bang:* que se produjo dentro del primer minuto mismo de la existencia del universo en expansión. No es necesario recalcar lo precario del *status* científico de esta especulación (debida originalmente a Gamow) y, puesto que tenemos que acudir a teorías de este tipo en nuestro intento de reducir la química a mecánica cuántica, difícilmente podrá pretenderse que esta reducción se ha realizado sin residuo.

La verdad es que hemos reducido la química, al menos en parte, a cosmología en vez de a teoría física. Es cosa admitida por todos que la cosmología relativista clásica moderna empezó con una teoría física aplicada; pero, como dice Hermann Bondi, esos tiempos han pasado ya y tenemos que hacer frente al hecho de que algunas de nuestras ideas (por ejemplo, las que empezaron con Dirac y Jordan) podrían describirse casi como intentos de reducir la teoría física a cosmogonía. Y tanto la cosmogonía como la cosmología, aunque son partes de la física enormemente fascinantes, y aunque están convirtiéndose en más fácilmente contrastables, siguen siendo casi casos límite de la ciencia física y apenas suficientemente maduros como para servir como bases para la reducción de la química a física. Esta es una de las razones por las que considero la llamada reducción de la química a física incompleta y algo problemática; aunque, por supuesto, me alegro enormemente de la aparición de todos esos nuevos problemas.

VII

Pero hay un segundo residuo de la reducción de la química a física. Nuestra concepción actual es que el hidrógeno solo, o más bien su núcleo, es el material de construcción de todos los núcleos restantes. Creemos que los núcleos positivos se repelen fuertemente entre sí eléctricamente hasta distancias muy cortas, pero

[16] Esta teoría puede estar ahora amenazada por la nueva teoría de las desviaciones del rojo, propuesta por J. C. Pecker, A. P. Roberts y J. P. Vigier, «Non-velocity redshifts and photon-photon interactions», *Nature*, 237, 1972, páginas 227-9.

que para distancias aún más cortas (que se consiguen sólo si la repulsión es vencida por tremendas presiones o velocidades) se atraen mutuamente por fuerzas nucleares.

Pero esto significa que atribuimos al núcleo de hidrógeno propiedades de relación que son inoperantes en la inmensa mayoría de las condiciones en las que los núcleos de hidrógeno existen en nuestro universo. Es decir, esas fuerzas nucleares son potencialidades que sólo entran en actividad tarde, bajo condiciones que son extremadamente raras: bajo temperaturas y presiones tremendas. Pero esto significa que la teoría de la evolución de la tabla periódica se parece mucho a una teoría de propiedades esenciales que tienen el carácter de *predestinación o de armonía preestablecida* [17]. En cualquier caso, un sistema solar como el nuestro depende, según las teorías actuales, de la preexistencia de esas propiedades o, más bien, potencialidades.

Además, la teoría del origen de los elementos más pesados en las explosiones de supernovas introduce un segundo tipo de *predestinación o armonía preestablecida.* Porque equivale a afirmar que las fuerzas gravitatorias (aparentemente, las más débiles de todas y, hasta el momento, sin relación con las fuerzas nucleares o electromagnéticas) [18] pueden, en grandes acumulaciones de hidrógeno, hacerse tan poderosas como para vencer la tremenda repulsión eléctrica entre los núcleos y hacerlos fundirse debido a la acción de fuerzas nucleares. Aquí, la armonía está entre las potencialidades inherentes a las fuerzas nucleares y a las de gravitación. No quiero decir que una filosofía de la armonía preestablecida tenga que ser falsa. Pero no creo que el recurso a una armonía preestablecida pueda considerarse una reducción satisfactoria; sugiero que ese recurso es una admisión del fracaso del método de reducir una cosa a otra.

Así pues, la reducción de la química a física está lejos de ser completa, aunque admitamos unos supuestos favorables poco realistas. Antes bien, esta reducción supone una teoría de la evolución cósmica o cosmogenia y además dos tipos de armonía preestablecida, para permitir que las potencialidades latentes, o las propensiones relativas de baja probabilidad inherentes en el áto-

[17] Utilizo aquí el término «armonía preestablecida» para resaltar que nuestra explicación no es en términos de las propiedades físicas manifiestas del átomo del hidrógeno. En cambio, una propiedad del núcleo de hidrógeno, desconocida e insospechada hasta el momento, fue postulada y utilizada como explicación.

[18] Véase la nota 15 más arriba.

mo de hidrógeno, sean activadas. Parece, sugiero yo, que deberíamos reconocer que estamos operando con las ideas de *emergencia* y de *propiedades emergentes* [19]. De este modo, vemos que esta interesantísima reducción nos ha dejado con una extraña visión del universo, extraña, en todo caso, para el reduccionista, que es lo que quería dejar patente en esta sección.

VIII

Resumiendo lo que se ha dicho hasta aquí: he tratado de poner en claro el problema de la reducción con ayuda de ejemplos y he tratado de mostrar que algunas de las reducciones más impresionantes en la historia de las ciencias físicas distan mucho del éxito completo y dejan un residuo. Podría sostenerse (pero véase la nota 5 más arriba) que la teoría de Newton fue una reducción de pleno éxito de las teorías de Kepler y Galileo. Pero incluso aunque supongamos que sabemos mucha más física de la que sabemos en realidad, y que tenemos una teoría de campo unificada que implica con gran aproximación la relatividad general, la teoría cuántica y los cuatro tipos de fuerzas como casos especiales, incluso entonces podemos decir que la química no se ha reducido a física sin residuo. En realidad, la llamada reducción de la química es a una física que presupone evolución, cosmología y cosmogenia, y la existencia de propiedades emergentes.

Por otro lado, en nuestras tentativas de reducción que no alcanzan pleno éxito, especialmente de la química a física, hemos

[19] Utilizo aquí el término «emergente» para indicar un paso evolutivo aparentemente imprevisible. Véase *El yo y su cerebro* [*S. & I. B. 1977*], capítulo P1, secciones 6-9. Sin querer propugnar lo que podría llamarse un programa de investigación anti-reduccionista para la biología, lo que sigue parece un comentario razonable de la situación. El programa mecanicista de investigación para la física se desbarató con el intento de incluir la electricidad y el magnetismo o, más precisamente, con la introducción por Faraday de las fuerzas no-centrales. (El intento de Maxwell de reducir esas fuerzas no-centrales a teoría newtoniana por medio de la construcción de un modelo mecánico del éter resultó sumamente fructífero al sugerirle sus ecuaciones de campo, pero no tuvo éxito y hubo de ser abandonado.) La comprensión por parte de Einstein de que las teorías de Newton y Maxwell no se funden llevó a la relatividad general. Así pues, los físicos tuvieron que aceptar una teoría radicalmente nueva en lugar de una reducción. Un destino similar le sobrevino a la física cuando tanto la teoría mecánica como la electromagnética en la forma unificada, que debemos a Lorentz y a Einstein, fueron aplicadas a problemas nuevos y en gran medida estadísticos de la microestructura de la materia. Esto llevó a la mecánica cuántica. No podemos descartar la posibilidad de que la inclusión de problemas biológicos lleve a una mayor expansión y revisión de la física.

aprendido una cantidad increíble de cosas. Los nuevos problemas han hecho surgir nuevas teorías conjeturales, y algunas de ellas, como la fusión nuclear, han llevado no sólo a experimentos que las han corroborado, sino también a una nueva tecnología. Así, desde el punto de vista del método, nuestros programas de reducción han llevado a grandes éxitos, aunque pueda decirse que las reducciones intentadas, como tales, han fracasado.

IX

La historia que se ha contado aquí y la lección que se desprende de ella difícilmente sosprenderán a un biólogo por inesperadas. También en biología, el reduccionismo (bajo la forma de fisicalismo o materialismo) ha tenido un éxito grande, aunque no total. Pero, incluso en los aspectos en los que no ha tenido éxito, ha llevado a nuevos problemas y a nuevas soluciones.

Quizá podría expresar mi opinión como sigue: como filosofía, el reduccionismo es un fracaso. Desde el punto de vista del método, los intentos de reducciones detalladas han llevado a un éxito asombroso tras otro y sus fracasos han sido sumamente provechosos para la ciencia.

Es quizá comprensible que algunos de los que han obtenido esos éxitos científicos no hayan percibido el fracaso de la filosofía. Quizá mi análisis del éxito y el fracaso del intento de reducir por completo la química a física cuántica les dé qué pensar y les haga volver a fijar su atención en el problema.

X

Lo hasta aquí expuesto puede considerarse una elaboración de la breve observación de Jacques Monod en el prefacio a *El azar y la necesidad* [20]:

> No se puede predecir ni resolver toda la química con la ayuda de la teoría cuántica [o reducirlo a teoría cuántica], que sin duda constituye, no obstante, la base universal.

---

[20] *El azar y la necesidad,* 1970, p. 10.

En el mismo libro, Monod hace también una sugerencia (desde luego, no una afirmación) respecto al origen de la vida, que es muy sorprendente y que podemos considerar desde el punto de vista que se ha alcanzado aquí. La sugerencia de Monod es que la vida emergió de la materia inanimada por una combinación sumamente improbable de circunstancias fortuitas y que esto puede no haber sido simplemente un suceso de baja probabilidad, sino de probabilidad cero; en realidad, un suceso *único*.

Esta sugerencia es contrastable experimentalmente (como ha señalado Monod en conversación con Eccles). Si consiguiéramos producir vida bajo ciertas condiciones experimentales bien definidas, entonces la hipótesis de la unicidad del origen de la vida quedaría refutada. Así pues, la sugerencia es una hipótesis científica contrastable, aunque no lo parezca a primera vista.

¿Qué otra cosa, además, hace plausible la sugerencia de Monod? Existe el hecho de la unicidad del código genético, pero eso podría ser, como señala Monod, el resultado de la selección natural. Lo que hace del origen de la vida, y del código genético, un enigma inquietante es esto: el código genético no tiene ninguna función biológica a menos que se traduzca, es decir, a menos que lleve a la síntesis de las proteínas cuya estructura está dispuesta por el código genético. Pero, como señala Monod, la maquinaria por la que la célula (al menos la célula no primitiva, que es la única que conocemos) traduce el código «consiste por lo menos en cincuenta componentes macromoleculares *que están, ellos mismos, codificados en el D. N. A.*» [21]. Así pues, el código no puede traducirse si no es usando ciertos productos de su traducción. Esto constituye un círculo verdaderamente desconcertante: un círculo vicioso, parece, para cualquier tentativa de formular un modelo, o una teoría, de la génesis del código genético.

Ocurre entonces que podemos encontrarnos con la posibilidad de que el origen de la vida (como el origen del universo) se convierte en una barrera impenetrable para la ciencia y en un residuo para todas las tentativas de reducir la biología a química y física. Porque incluso aunque la sugerencia de Monod de la unicidad del origen de la vida fuera refutable —por medio de tentativas de reducción, naturalmente— equivaldría, si fuese cierta, a negar la posibilidad de reducciones de pleno éxito. Con esta sugerencia, Monod, que es un reduccionista por razones de método, llega a

[21] *Ibid*, p. 143.

la posición que es, yo creo, a la que todos nos vemos forzados en vista de nuestra discusión anterior de la reducción de la química a física. Es la posición de un reduccionista crítico que continúa intentando reducciones aunque desespere del éxito final. Sin embargo, es en la continuación de los intentos de reducción, como resalta Monod en otra parte de su libro, y no en una sustitución de los métodos reduccionistas por otros «holistas», en donde reside nuestra principal esperanza: nuestra esperanza de aprender más sobre viejos problemas, y de descubrir nuevos problemas que, a su vez, nos lleven a nuevas soluciones, a nuevos descubrimientos.

No quiero tratar aquí del holismo con ningún detalle, pero pueden resultar necesarias unas palabras. El uso de métodos experimentales holistas (tales como los trasplantes de células en embriones), aunque están inspirados por el pensamiento holista, pueden muy bien considerarse metodológicamente reduccionistas. Las teorías holistas son, por otro lado, trivialmente necesarias en la descripción, incluso de un átomo o molécula, por no hablar de un organismo o una población genética. No hay límite a la variedad de conjeturas posiblemente fructíferas, sean holistas o no [22]. En vista de mi tesis principal, sólo surgen dudas respecto al carácter de los métodos experimentales en biología: si no son todos, en mayor o menor grado, de un carácter reduccionista. (Una situación similar surge, incidentalmente, como me recuerda David Miller, con respecto a las teorías deterministas o indeterministas. Aunque, como he argüido en el núcleo de este tomo, tenemos que ser metafísicamente *indeterministas*, metodológicamente deberíamos continuar la búsqueda de leyes deterministas o causales, excepto allí donde los problemas a resolver tengan, ellos mismos, un carácter probabilista.)

Incluso aunque la sugerencia de Monod sobre la unicidad del origen de la vida fuese refutada por la creación de vida a partir de materia inanimada bajo condiciones claramente repetibles, esto no equivaldría a una reducción completa. No quiero argüir *a priori* que una reducción sea imposible; pero hemos creado vida a partir de la vida durante mucho tiempo sin entender lo que está-

[22] Antes de que podamos siquiera intentar una reducción, necesitamos un conocimiento tan extenso y tan detallado como sea posible de lo que intentemos reducir. Así, antes de poder intentar una reducción tenemos que trabajar en el nivel de la cosa a reducir; es decir, en el nivel de las «totalidades». Véase mi *Conocimiento objetivo* [*Obj. Kn.*].

bamos haciendo y antes de tener siquiera la menor idea sobre biología molecular o sobre el código genético. Así pues, es ciertamente posible que produzcamos vida a partir de materia inanimada sin una plena comprensión físico-química de lo que estamos haciendo; por ejemplo, de cómo conseguimos romper el círculo vicioso inherente a la traducción del código.

En cualquier caso, podemos decir que el inopinado descubrimiento de la biología molecular ha hecho del problema del origen de la vida un enigma mayor de lo que ya era: hemos adquirido nuevos y más profundos problemas.

XI

La tentativa de reducir la química a física exige, pues, la introducción en la física de una teoría de la evolución; es decir, un recurso a la historia de nuestro cosmos. Una teoría de la evolución es, parece, aún más indispensable en biología. Y lo mismo ocurre, además, con la idea de propósito o teleología o (por usar el término de Mayr) teleonomía, o la idea, muy similar, de resolución de problemas; una idea que es bastante ajena al tema de las ciencias no-biológicas (aunque el papel que juegan en esas ciencias los máximos y los mínimos y el cálculo de variaciones ha sido considerado como remotamente análogo).

El gran logro de Darwin fue, naturalmente, demostrar que es posible explicar teleología en términos no teleológicos o términos causales corrientes. El darwinismo es la mejor explicación que tenemos. En este momento no existen hipótesis que le hagan una competencia seria [23].

Los problemas y su resolución parecen emerger con la vida. Aunque hay algo parecido a la selección natural operando antes del origen de la vida —por ejemplo, una selección de los elementos más estables debido a la desintegración radiactiva de los menos estables—, no podemos decir que, para los núcleos atómicos, la superviviencia sea un «problema» en ningún sentido de la palabra. Y la estrecha analogía entre cristales y micro-organismos y sus partes moleculares (organelas) termina también aquí. Los cristales no tienen problemas de crecimiento o de propagación o de supervivencia. Pero la vida se enfrenta a los problemas de la su-

[23] Véase *Conocimiento objetivo* [*Obj. Kn.*], capítulos 6 y 7.

pervivencia desde el principio mismo. En realidad, podemos describir la vida, si queremos, como resolución de problemas y a los organismos vivos como los únicos conjuntos del universo capaces de resolver problemas (las computadoras son *instrumentos para* la resolución de problemas, pero no quienes los resuelvan, en este sentido).

Esto no significa que tengamos que atribuir a toda forma de vida una *conciencia* de los problemas a resolver. Incluso en el nivel humano resolvemos constantemente muchos problemas, tales como el de conservar el equilibrio, sin percatarnos de ello.

XII

Poca duda cabe de que los animales poseen conciencia y de que, a veces, pueden ser, incluso, conscientes de un problema. Pero la emergencia de la conciencia en el reino animal es, quizá, un misterio tan grande como el origen de la vida.

Sobre esto sólo quiero decir que el pansiquismo, o hilozoísmo, o la tesis de que la materia, en general, está dotada de conciencia (de bajo grado), no me parece ayudar en lo más mínimo. Es, si es que se toma en serio, una teoría distinta de la predestinación, de una armonía preestablecida. (Era, desde luego, parte de la forma original de la teoría de Leibniz sobre la armonía preestablecida.) Porque, en la materia sin vida, la conciencia no tiene función ninguna; y, si (con Leibniz, Diderot, Buffon, Haeckel y muchos otros) atribuimos conciencia de las partículas no-vivientes (mónadas, átomos), lo hacemos con la vana esperanza de que eso nos ayude a explicar la presencia de aquellas formas de conciencia que tienen alguna función en los animales.

Porque no cabe duda de que la conciencia cumple una función en los animales y puede contemplarse como si fuera un órgano corporal. Tenemos que suponer, por difícil que sea, que es un producto de la evolución, de la selección natural.

Aunque esta observación puede constituir un programa para una reducción, no es todavía una reducción, y la situación del reduccionista parece algo desesperada, lo que explica por qué los reduccionistas han adoptado la hipótesis del pansiquismo o por qué, más recientemente, han negado totalmente la existencia de la conciencia (la conciencia, digamos, de un dolor de muelas).

A pesar de que esta filosofía conductista está muy en boga ac-

tualmente, una teoría de la inexistencia de la conciencia no puede tomarse más en serio que, sugiero, una teoría de la inexistencia de la materia. Las dos teorías «resuelven» el problema de la relación entre el cuerpo y la mente. La solución es, en ambos casos, una simplificación radical: es la negación del cuerpo o la de la mente. Pero, en mi opinión, es demasiado fácil [24]. Diré algo más sobre esta «cuestión pendiente» y especialmente sobre el pansiquismo, más adelante, en la sección XIV, donde critico el paralelismo psicofísico.

XIII

De las tres «cuestiones pendientes de la reducción» mencionadas al principio de este artículo he hablado brevemente de dos. Ahora trataré de la tercera, la cuestión de la reducción de la conciencia humana del yo y de la creatividad de la mente humana.

Como a menudo ha resaltado Sir John Eccles, esta tercera cuestión constituye el problema de «la vinculación mente-cerebro»; y Jacques Monod llama al problema del sistema nervioso central humano «la segunda frontera», comparando su dificultad con la «primera frontera», el problema del origen de la vida.

Sin duda, esta frontera es una región peligrosa de habitar, especialmente para el biólogo profano; no obstante, diré que los intentos de una reducción parcial me parecen más esperanzadores en esta región que en la de la segunda cuestión. Como en la región de la primera cuestión, me parece que pueden descubrirse aquí más problemas nuevos con métodos reduccionistas, y quizá incluso resolverlos, que en la región de la segunda cuestión; una región que me parece a mí estéril en comparación. Apenas necesito resaltar que una reducción de un éxito completo en cualquiera de las tres regiones me parece muy improbable, si no imposible.

Con esto, puede decirse, quizá, que he cumplido mi promesa de discutir, o en cualquier caso mencionar, estas tres cuestiones pendientes de la reducción que enumeré al principio de este ensayo. Pero quiero decir algo más sobre la tercera de ellas —sobre el problema cuerpo-mente, o mente-cuerpo— antes de seguir con mi tesis de la incompletabilidad de toda ciencia.

---

[24] Véanse *Conocimiento objetivo* [*Obj. Kn.*], capítulo 8, y *El yo y su cerebro* [*S. & I. B.*], esp. capítulo P3, sección 19.

Considero el problema de la emergencia de la conciencia en los animales (cuestión 2), de comprenderla y, quizá, de reducirla a fisiología, como muy probablemente insoluble; y pienso lo mismo sobre el problema ulterior de la emergencia de la conciencia específicamente humana del yo (cuestión 3); es decir, el problema cuerpo-mente. Pero yo creo que podemos arrojar al menos cierta luz sobre el problema del yo humano.

Yo soy, en muchos aspectos, dualista cartesiano [25], aunque prefiero describirme a mí mismo como pluralista, y, desde luego, no estoy de acuerdo con ninguna de las dos sustancias de Descartes. La materia, como hemos visto, no es una sustancia elemental con la propiedad esencial de extensión, sino que consiste en estructuras complejas sobre cuya constitución sabemos mucho, incluida la explicación de la mayor parte de su «extensión»: ocupa mucho espacio (si no todo) por medio de la repulsión eléctrica.

Mi primera tesis es que la conciencia humana del yo, con su unidad aparentemente irreducible, es enormemente compleja y que puede ser, quizá, explicable en parte.

He sugerido [26] que la conciencia humana superior, o conciencia de la propia individualidad, está ausente en los animales. También he sugerido que la conjetura de Descartes que localiza el alma humana en la glándula pineal puede ser tan absurda como a menudo se la ha hecho parecer, y que, en vista de los resultados de Sperry con hemisferios cerebrales divididos [27], puede que haya que buscar la localización en el centro del habla, en el hemisferio izquierdo del cerebro. Como Eccles me ha informado más recientemente [28], los experimentos posteriores de Sperry apoyan esta suposición hasta cierto punto: el hemisferio derecho puede describirse como el de un animal muy inteligente, mientras que sólo el hemisferio izquierdo parece ser humano y consciente del propio yo.

Yo había basado mi suposición en el papel que atribuyo al desarrollo de un lenguaje específicamente humano.

Todo lenguaje animal —de hecho, casi toda la conducta ani-

---

[25] Véanse *El desarrollo del conocimiento científico* [*C. & R.*], capítulo 12; y *El yo y su cerebro* [*S. & I. B.*], capítulo P4.

[26] En un ciclo de conferencias que tuvo lugar en la Emory University, en mayo de 1969, y durante muchos años en mis conferencias en la London School of Economics. Véase también *El yo y su cerebro* [*S. & I. B.*].

[27] *El yo y su cerebro* [*S. & I. B.*], especialmente los capítulos E4 y E5.

[28] J. C. Eccles, «Unconscious actions emanating from the human cerebral cortex», inédito, 1972.

mal— tiene una función *expresiva* (o sintomática) y una función *comunicativa* (o señaladora), como lo ha indicado Karl Bühler [29]. Pero el lenguaje humano tiene además algunas funciones más, que son características de él y le convierten en un «lenguaje» en el sentido más estricto y más importante de la palabra. Bühler llamó la atención sobre la función básica *descriptiva* del lenguaje humano, y señaló más tarde [30] que hay aún más funciones (tales como la preceptiva, la consultiva, etc.), de las cuales la más importante y característica de los seres humanos es la función *argumentativa.* (Alf Ross señala que pueden añadirse muchas otras funciones; por ejemplo, las de dar órdenes o hacer ruegos o promesas) [31].

Nunca he creído que ninguna de esas funciones fuese reducible a ninguna de las otras, y, menos que ninguna, las dos funciones superiores (descripción y argumentación) a las dos inferiores (expresión y comunicación). Estas, dicho sea de paso, siempre están presentes, lo que quizá explique por qué tantos filósofos las confunden con las propiedades que son características del lenguaje humano.

Mi tesis es que, con las funciones superiores del lenguaje humano, emerge un nuevo mundo: el mundo de los productos de la mente humana. Lo he llamado «Mundo 3». Ya he descrito cómo uso ese término —y también «Mundo 1» y «Mundo 2»— en el epílogo de este volumen. Brevemente, llamo al mundo de la materia física, de los campos de fuerza, etc., «Mundo 1»; al mundo de la conciencia y quizá también de la experiencia subconsciente, «Mundo 2»; y «Mundo 3» especialmente al mundo del lenguaje hablado (o escrito o impreso), como el relato de historias, la creación de mitos, teorías, problemas teóricos, detección de errores y argumentación (los mundos de los productos artísticos y de las instituciones sociales pueden, o bien ser incluidos en el «Mundo 3» o ser llamados «Mundo 4» y «Mundo 5»: eso es una cuestión de gustos o conveniencias).

Introduje esos términos para hacer hincapié en la *autonomía* (limitada) de esas regiones. La mayoría de los materialistas o fisicalistas o reduccionistas afirman que, de esos tres mundos, sólo existe realmente el «Mundo 1» y que es, por tanto, autónomo. Sustituyen el «Mundo 2» por la conducta y el «Mundo 3», parti-

[29] Karl Bühler, *Sprachtheorie*, 1934.

[30] Véase *El desarrollo del conocimiento científico* [*C. & R.*], capítulos 4 y 12.

[31] Alf Ross, «The rise and fall of the doctrine of perfomatives», en *Contemporary Philosophy in Scandinavia*, ed. R. E. Olsen y A. M. Paul, 1972, pp. 197-212.

cularmente, por la conducta verbal. (Como se ha indicado más arriba, ésta es una de esas maneras, demasiado fáciles, de resolver el problema cuerpo-mente: la de negar la existencia de la mente humana y de la conciencia humana del propio yo; es decir, de las cosas que yo considero entre las más notables y asombrosas del universo; la otra forma igualmente fácil es la del inmaterialismo de Berkeley y de Mach.) La tesis de que sólo existen las sensaciones y que la materia no es más que un «constructo» de sensaciones.

XIV

Hay, principalmente, cuatro posturas respecto a la interrelación entre el cuerpo, o el cerebro, y la mente.

1) Negación de la existencia del «Mundo 1» de los estados físicos, es decir, inmaterialismo, como el sustentado por Berkeley y Mach.

2) Negación del «Mundo 2» de los estados o sucesos mentales, una concepción común entre ciertos materialistas, fisicalistas y conductistas filosóficos, o filósofos que mantienen la identidad entre el cerebro y la mente.

3) Afirmación de un paralelismo absoluto entre los estados mentales y los estados del cerebro; una posición que recibe el nombre de «paralelismo psicofísico». Fue introducida por primera vez en la escuela cartesiana por Geulincx, Spinoza, Malebranche y Leibniz, principalmente, para evitar ciertas dificultades de la concepción cartesiana. (Como el epifenomenalismo, priva a la conciencia de funciones biológicas.)

4) Afirmación de que los estados mentales pueden interactuar con los estados físicos. Esta era la concepción de Descartes que muchos consideran que fue reemplazada por 3).

Mi propia postura es que un paralelismo entre el cerebro-mente casi no tiene más remedio que existir, *hasta cierto punto.* Ciertos reflejos, tales como parpadear cuando se ve un objeto acercándose rápidamente, tienen toda la apariencia de poseer un carácter más o menos paralelista: la reacción muscular (en la que, sin duda, está implicado el sistema nervioso central) se repite con regularidad cuando se repite la impresión visual. Si se nos llama la atención hacia este reflejo, podemos ser conscientes de que

ocurre, y lo mismo pasa con algunos de los reflejos restantes (pero, desde luego, no con todos).

No obstante, yo creo que la tesis de un paralelismo psicofísico *completo* —postura 3)— es errónea, incluso en algunos casos en los que los implicados son simples reflejos. *Así pues, yo propongo una forma de interaccionismo psicofísico.* Esto supone (como lo vio Descartes) *la tesis de que el «Mundo 1» físico no está casualmente cerrado,* sino abierto al «Mundo 2», abierto a los estados mentales y a los sucesos mentales. Esta es, quizá, una tesis poco atractiva para el físico, pero creo que está apoyada por el hecho de que el «Mundo 3» (incluidas sus regiones autónomas) actúa sobre el «Mundo 1» a través del «Mundo 2». (Mucho de esto se ha expuesto en el epílogo de este libro.)

Estoy dispuesto a aceptar la concepción según la cual, cuando ocurre algo en el «Mundo 2», algo relacionado con ello ocurre en el «Mundo 1» (en el cerebro). Pero para hablar de un paralelismo completo y cabal tendríamos que afirmar que «el mismo» estado o suceso mental va siempre acompañado por un estado fisiológico exactamente correspondiente, y viceversa.

Como se ha indicado, estoy de acuerdo en admitir que hay algo correcto en esa afirmación y que, por ejemplo, la estimulación eléctrica de ciertas regiones del cerebro puede dar lugar regularmente a ciertos movimientos o sensaciones característicos. Pero yo pregunto si, como regla universal de los estados mentales, tiene la aserción algún contenido o si es una aserción vacía. Porque podemos tener un paralelismo entre los elementos del «Mundo 2» y los procesos del cerebro, o entre *Gestalten* del «Mundo 2» y procesos del cerebro, pero difícilmente podemos hablar de un paralelismo entre un proceso enormemente complicado, único e inanalizable del «Mundo 2» y un proceso del cerebro. Y hay muchos menos sucesos del «Mundo 2» en nuestras vidas que son únicos. Incluso aunque no tengamos en cuenta la novedad creativa, oír una melodía dos veces y reconocer que es la misma no es una repetición del mismo suceso del «Mundo 2», simplemente porque la segunda audición de la melodía está relacionada con el acto de *reconocer* la melodía, que estaba ausente la primera vez. El objeto del «Mundo 1» (en este caso un objeto que pertenece también al «Mundo 3») es el que se repite, no el suceso del «Mundo 2». Sólo si pudiéramos aceptar una especie de teoría del «Mundo 2» que, como la psicología asociacionista, considerase que los sucesos del «Mundo 2» están compuestos de elementos recurrentes, podría-

mos hacer una clara distinción entre la experiencia repetida del «Mundo 2» —el *oír* la misma melodía— y la parte no repetida, el *reconocimiento* de que es la misma melodía (donde la experiencia del reconocimiento es, a su vez, capaz de recurrir en otros contextos). Pero creo que está claro que una psicología tan atomista y analítica es bastante incapaz de llevarnos lejos.

El «Mundo 2» es muy complejo. Si nos ocupamos sólo de campos como la percepción sensorial (es decir, la percepción de objetos del «Mundo 1») podemos creer que el «Mundo 2» puede analizarse por métodos atómicos o moleculares, por ejemplo, métodos *Gestalt* (métodos que, creo yo, son todos ellos poco provechosos comparados con los métodos biológicos o funcionales de Egon Brunswik). Pero la aplicación de esos métodos resulta bastante inadecuada si consideramos nuestros intentos, únicos, de inventar o entender un objeto del «Mundo 3», como un problema o una teoría.

La forma en que nuestro pensamiento o nuestro entendimiento interactúan con los intentos de formulación lingüística y son influenciados por ella; la forma en que tenemos un vago sentimiento respecto de un problema o de una teoría, que se clarifica cuando tratamos de formularlos y queda aún más claro cuando lo escribimos y somos críticos respecto a nuestros intentos de resolverlo; la forma en que un problema puede cambiar y seguir siendo, en cierto sentido, el mismo problema; la forma en que una sucesión de pensamientos está conectada por un lado y articulada por otro: todo esto me parece estar más allá de métodos atomistas o analíticos, incluidos los interesantes métodos moleculares de la psicología de la *Gestalt*. (Por cierto que, mientras que los psicólogos de la *Gestalt* enseñan que las hipótesis son *Gestalten*, yo enseño que las *Gestalten* son hipótesis: interpretaciones, de las señales recibidas.)

Además, tenemos razones para pensar que, a menudo, si una región del cerebro se destruye, otra región se hace cargo de sus funciones, con muy poca o quizá sin ninguna interferencia con el «Mundo 2». Otro argumento en contra del paralelismo, y esta vez basado en experimentos del «Mundo 1» en vez de en la consideración, necesariamente vaga, de las experiencias más complejas del «Mundo 2».

Todo esto suena, desde luego, muy anti-reduccionista; y como filósofo que mira a este mundo nuestro, con nosotros en él, yo realmente desespero de encontrar una reducción final. Pero,

metodológicamente, eso no me lleva a un programa de investigación anti-reduccionista. Sólo lleva a la predicción de que, con el aumento de nuestras tentativas de reducción, se expandirá nuestro saber y, con él, la serie de los problemas sin resolver.

XV

Volvamos ahora al problema de la conciencia específicamente humana del yo; mi sugerencia es que emerge a través de la interacción (realimentación, si se quiere) entre el «Mundo 2» y los «Mundos 1 y 3». Mis argumentos en relación con el papel que juega el «Mundo 3» son los que siguen:

La conciencia humana del propio yo está basada, entre otras cosas, en una serie de *teorías* sumamente abstractas. Los animales y las plantas tienen, sin duda, un sentido del tiempo y expectativas temporales. Pero hace falta una *teoría* casi explícita del tiempo (*pace* Benjamin Lee Whorf) para verse a sí mismo en posesión de un pasado, un presente y un futuro; en posesión de una historia personal; y consciente de la propia identidad personal (ligada a la identidad del propio cuerpo) a lo largo de esa historia. Así pues, es una *teoría* que, durante los períodos de sueño, cuando perdemos la continuidad de la conciencia, nosotros —nuestros cuerpos— continuamos esencialmente iguales; y sobre la base de esta teoría es como podemos recordar conscientemente acontecimientos pasados (en lugar de estar simplemente influidos por ellos en nuestras expectaciones y reacciones que, sugiero yo, son la forma más primitiva que adopta la memoria en los animales).

No hay duda de que algunos animales tienen personalidad; tienen algo muy análogo al orgullo y la ambición, y aprenden a responder a un nombre. Pero la conciencia humana del yo está anclada en el lenguaje y (tanto explícita como implícitamente) en teorías formuladas. Un niño aprende a usar el nombre de sí mismo y finalmente un nombre como «ego» o «yo» y aprende a usarlo con la conciencia de la continuidad de su cuerpo, y de sí mismo; y también lo combina con el conocimiento de que la conciencia no es siempre ininterrumpida. La gran complejidad y el carácter no sustancial del alma humana, o del yo humano, quedan claros si recordamos que hay casos de hombres que han olvidado quiénes eran; han olvidado parte o toda su historia pasada, pero han retenido, o quizá recobrado, al menos parte de su identi-

dad. En cierto sentido, su memoria no se ha perdido, porque *recuerdan cómo* andar, comer e incluso hablar. Pero no *recuerdan que* ellos vienen de, digamos, Bristol o cuáles sean sus nombres y direcciones. En la medida en que no encuentran el camino de sus casas (lo que los animales suelen hacer) su conciencia del propio yo está afectada incluso más allá del nivel normal de la memoria animal. Pero, si no han perdido la capacidad de hablar, alguna conciencia humana les queda, que va más allá de la memoria animal.

No soy un gran amigo del psicoanálisis, pero sus hallazgos parecen apoyar la idea de la complejidad del yo humano, en contraposición con cualquier recurso cartesiano a una sustancia pensante. Mi principal argumento es que la conciencia del yo humano implica, como mínimo, un percatarse de la (sumamente teórica) continuidad temporal o histórica del propio cuerpo; un percatarse de la conexión entre la propia memoria consciente y el cuerpo singular, único, que es el propio; y la conciencia de la interrupción normal y periódica de la propia conciencia por el sueño (que, una vez más, implica una teoría del tiempo y de la periodicidad temporal). Además, implica la conciencia de pertenecer, local y socialmente, a un lugar determinado y a un determinado grupo de gente. Sin duda, mucho de esto tiene una base instintiva y lo comparten los animales. Mi tesis es que incluso en elevarlo al plano de la conciencia humana inexpresa el lenguaje humano o la interacción entre los «Mundos 2 y 3» juegan un papel importante.

Está claro que la unidad del yo humano se debe en gran parte a la memoria, y la memoria puede atribuírseles no sólo a los animales, sino también a las plantas (y quizá incluso, en cierto sentido, a estructuras no orgánicas, como los magnetos). Es, por tanto, de la mayor importancia comprender que el recurso a la memoria como tal no es suficiente para explicar la unidad del yo humano. Lo que se necesita no es tanto la memoria humana «ordinaria» (de los acontecimientos pasados), sino una memoria de teorías que relacionen la conciencia de tener un cuerpo con teorías del «Mundo 3» sobre los cuerpos (es decir, con la física), una memoria que es de la naturaleza de una «captación» de las teorías del «Mundo 3». Incluye las disposiciones que nos permiten echar mano de teorías explícitas del «Mundo 3» cuando lo necesitamos, con la sensación de que poseemos esas disposiciones y de que podemos hacer uso de ellas para articular esas teorías si lo necesitamos. (Esto, desde luego, explicaría hasta cierto punto la diferencia

entre la conciencia humana del yo con su dependencia del lenguaje humano y la conciencia animal.)

XVI

Me parece que estos hechos establecen la imposibilidad de cualquier reducción del «Mundo 2» humano, el mundo de la conciencia humana, al «Mundo 1» humano; es decir, esencialmente, a la fisiología del cerebro. Porque el «Mundo 3» es, al menos en parte, autónomo e independiente de los otros dos mundos. Si la parte autónoma del «Mundo 3» puede interactuar en el «Mundo 2», entonces el «Mundo 2», o al menos eso me parece, no puede ser reducible al «Mundo 1».

Mis ejemplos habituales de la autonomía parcial del «Mundo 3» están tomados de la aritmética.

Como ya he explicado en el epílogo a este libro, la serie infinita de los números naturales proporciona ese ejemplo. Es una invención, un producto de la mente humana y una parte del lenguaje humano desarrollado. Existen, al parecer, lenguajes primitivos en los que uno sólo puede contar «uno, dos, muchos» y otros en los que sólo se puede contar hasta cinco. Pero, una vez que se ha inventado un método para contar sin límites, surgen autónomamente los problemas y las distinciones: los números pares e impares no son *inventados,* sino *descubiertos* en la serie de los números naturales, y lo mismo ocurre con los números primos y con los muchos problemas, resueltos y sin resolver, relacionados con ellos.

Esos problemas y los teoremas que los resuelven (tales como el teorema de Euclides, de que no existe un número primo máximo) surgen de forma autónoma; surgen como parte de la estructura interna de la serie, creada por el hombre, de los números naturales, e independientemente de lo que pensemos o dejemos de pensar. Pero podemos *captar* o *comprender* o *descubrir* esos problemas y resolver algunos de ellos. Así, nuestro pensamiento, que pertenece al «Mundo 2», depende en parte de los problemas autónomos y de la verdad objetiva de los teoremas, que pertenecen al «Mundo 3»; el «Mundo 2» no sólo crea al «Mundo 3», es en parte creado por el «Mundo 3» en una especie de proceso de realimentación.

Mi argumento, pues, discurre como sigue: el «Mundo 3», en

especial su parte autónoma, es claramente irreducible al «Mundo 1» físico. Pero, como el «Mundo 2» depende en parte del «Mundo 3», también es irreducible al «Mundo 1».

Los fisicalistas o reduccionistas filosóficos, como les he llamado, se ven así reducidos a negar la existencia de los «Mundos 1 y 2». Pero, con ello, la totalidad de la tecnología humana (sobre todo la existencia de computadoras) que tanto uso hace de los teoremas del «Mundo 3» se convierte en incomprensible; y tenemos que suponer que cambios tan violentos en el «Mundo 1» como los producidos por los constructores de aeropuertos o rascacielos se producen, en último término, por el propio «Mundo 1», sin la invención de las teorías del «Mundo 3» o los planes del «Mundo 2» que se basan en ellas: son cambios predestinados; son parte de una armonía preestablecida inherente, en último término, en los núcleos de hidrógeno.

Esos resultados me parecen absurdos; y el conductismo filosófico o fisicalismo (o la filosofía de la identidad entre la mente y el cuerpo) me parece que queda reducido a este absurdo.

XVII

El reduccionismo filosófico es, creo yo, un error. Se debe al deseo de reducir todo a una explicación última en términos de esencias y sustancias; es decir, a una explicación que no necesita explicaciones ulteriores, ni es capaz de darlas.

Una vez que renunciamos a la teoría de la explicación última, nos damos cuenta de que siempre podemos continuar preguntando «¿por qué?». Las preguntas de «por qué» nunca llevan a una respuesta última. Los niños inteligentes parecen saberlo, pero ceden ante los adultos que, desde luego, no pueden tener tiempo suficiente para lo que es, en principio, una serie de preguntas sin fin.

XVIII

Los «Mundos 1, 2 y 3», aunque autónomos en parte, pertenecen al mismo universo; interactúan. Pero puede mostrarse fácilmente que el conocimiento sobre el universo, si este conocimiento forma parte, él mismo, del universo, debe ser imposible de completar.

En el epílogo he mencionado el ejemplo del hombre que dibuja un mapa detallado de la habitación en la que trabaja, y los problemas que surgen cuando empieza a meter en su dibujo el mapa que está dibujando. Está claro que no puede completar su tarea.

Este y otros ejemplos ayudan a comprender por qué toda ciencia explicativa es incompletable; porque para ser completa tendría que dar una explicación de sí misma.

Implícito en el famoso teorema de Gödel, sobre la incompletatibilidad de la aritmética formalizada, hay un resultado más fuerte aún (pese a que usar el teorema de Gödel y otros teoremas de la incompletud de la matemática en este contexto equivale a usar artillería pesada contra una posición relativamente débil). Puesto que toda ciencia física usa aritmética (y puesto que, para un reduccionista, sólo la ciencia formulada en símbolos físicos tiene realidad), el teorema de la incompletud de Gödel hace incompleta a toda la ciencia física, lo que, para el reduccionista, mostraría que toda la ciencia es incompleta. Para el no-reduccionista, que no cree que toda la ciencia pueda ser reducida a ciencia formulada físicamente, la ciencia es incompleta de todas formas.

No sólo es erróneo el reduccionismo filosófico, sino que la creencia de que el método de reducción puede lograr reducciones completas también parece ser errónea. Vivimos en un mundo de evolución emergente, de problemas cuyas soluciones, si son resueltas, engendran nuevos y más profundos problemas. Así pues, vivimos en un mundo de emergente novedad; una novedad que, no general, no es completamente reducible a ninguno de los estudios precedentes.

No obstante, el método de intentar reducciones es muy fructífero, no sólo porque aprendemos mucho con sus éxitos parciales, por las reducciones parciales, sino también porque aprendemos de nuestros fracasos parciales, de los nuevos problemas que revelan nuestros fracasos. Los problemas abiertos son casi tan interesantes como sus soluciones; de hecho, serían igual de interesantes sino fuera porque casi toda solución abre, a su vez, un nuevo mundo de problemas abiertos.

MAS OBSERVACIONES SOBRE LA REDUCCION, 1981

I

Cuando escribí el *Addendum* anterior, a principios de 1972, me interesaba resaltar dos cosas. Una es el valor del intento de reducir: el frecuentemente increíble éxito de dichos intentos y la nueva comprensión que generan; la otra era que, a pesar de todo, no tenemos reducciones de un éxito completo, cuando «éxito» significa algo más que el aumento de nuestra clarividencia, de nuestra comprensión: cuando significa que se ha mostrado que una rama de nuestro saber, como la química, es completamente derivable de otra rama del saber, como la teoría atómica.

Al suscitar una seria duda de que existan tales reducciones totales, quería combatir lo que he llamado «reduccionismo filosófico», la anticipación algo dogmática de que las reducciones, tarde o temprano, tendrán pleno éxito, por alguna razón filosófica; dicho de otro modo, que tendrán éxito porque sabemos bastante sobre el mundo o sobre nosotros mismos o sobre el lenguaje o sobre la ciencia o sobre la filosofía o sobre qué se yo qué, para saber que el reduccionismo es verdadero.

A aquellos que dicen eso yo les respondo que no sabemos nada por el estilo y que el mundo es mucho más interesante y emocionante de lo que la filosofía reduccionista imagina.

II

Todavía recuerdo con vividez la emoción del descubrimiento del elemento 72 (hafnio), en 1922, como consecuencia de la maravillosa teoría cuántica de Niels Böhr sobre el sistema periódico de los elementos. Entonces nos pareció que había llegado el gran momento en que la química había sido reducida a teoría atómica; y fue, todavía me siento inclinado a decirlo, el momento más grandioso de todas las aventuras reduccionistas del siglo XX, sólo

superado quizá por el hallazgo que representó el descubrimiento de la estructura del D. N. A. por Click y Watson. Todavía tengo un libro de texto de 1929 en el que el progreso está espectacularmente representado por dos diagramas que reproduzco aquí (en recuerdo agradecido de su autor, Arthur Haas, y de mi amigo Franz Urbach, que le ayudó a completarlos) [1].

La teoría de Böhr llevó no sólo a la predicción de las propiedades químicas de los elementos, y de ahí a la predicción de las propiedades del todavía desconocido elemento 72 y así a su descubrimiento, sino que también permitió la predicción de algunas de sus propiedades ópticas; e incluso llevó a la predicción de algunas de las propiedades de los compuestos químicos.

Fue un gran momento en la historia de la materia. Sentimos, con razón, que *esto era lo que buscábamos:* Bohr había tocado el fondo rocoso. Y, sin embargo, un tipo de problema bastante diferente se cernía ya sobre el horizonte, iniciado por una sugerencia de Soddy (1910) y un descubrimiento de J. J. Thomsom (1913, año del modelo atómico de Böhr), y por la espectrocopia de masa

	I a	I b	II a	II b	III a	III b	IV a	IV b	V a	V b	VI a	VI b	VII a	VII b	VIII a	VIII b
1	1 H															2 He
2	3 Li		4 Be		5 B		6 C		7 N		8 O		9 F			10 Ne
3	11 Na		12 Mg		13 Al		14 Si		15 P		16 S		17 Cl			18 Ar
4	19 K	29 Cu	20 Ca	30 Zn	21 Sc	31 Ga	22 Ti	32 Ge	23 V	33 As	24 Cr	34 Se	25 Mn	35 Br	26 Fe 27 Co 28 Ni	36 Kr
5	37 Rb	47 Ag	38 Sr	48 Cd	39 Y	49 In	40 Zr	50 Sn	41 Nb	51 Sb	42 Mo	52 Te	43 Ma	53 J	44 Ru 45 Rh 46 Pd	54 X
6	55 Cs	79 Au	56 Ba	80 Hg	57—71 seltene Erden	81 Tl	72 Hf	82 Pb	73 Ta	83 Bi	74 W	84 Po	75 Re	86	76 Os 77 Ir 78 Pt	86 Em
7	87		88 Ra		89 Ac		90 Th		91 Pa		92 U					

TABLA 1

El sistema periódico de los elementos químicos, ordenados según su peso atómico y según fenómenos químicos y de otro tipo.

[1] Arthur Hass, *Atomtheorie* (Berlín y Leipzig, 1929), pp. 35 y 111. [Véase la discusión de Popper sobre Urbach en *Búsqueda sin término (U. Q.) Ed.*]

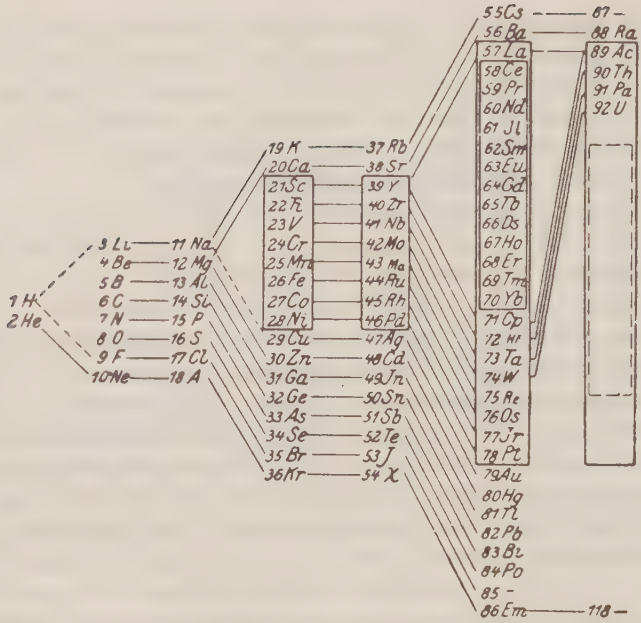

TABLA 2
El sistema periódico según la teoría de Niels Böhr sobre la estructura cortical de los átomos.

de F. W. Aston (1919). Y entonces llegó el zambombazo de Uray, el descubrimiento del agua pesada, que significaba que todas las medidas básicas de la química, las medidas de los pesos atómicos —el fenómeno básico de la química y del sistema periódico— estaban ligeramente equivocadas y tenían que ser revisadas.

Así, el fondo rocoso cedió de repente: de alguna manera, Niels Böhr había construido sobre una ciénaga. Pero su edificio seguía en pie.

Entonces llegó la mecánica cuántica y la teoría de London y Heitler. Y quedó bastante claro que la reducción de la química a física era una reducción sólo en principio; y que cualquier cosa parecida a una reducción completa estaba ahora a mayor distancia de lo que había parecido en 1922, el año del gran hallazgo.

Esto no es más que un vistazo a la historia, esbozado para hacer las cosas menos abstractas, porque ahora llego a un capítulo un tanto abstracto: a la lógica del reduccionismo.

III

Peter Madawar trata críticamente la reducción [2], usando la siguiente Tabla 3:

(4) Ecología/Sociología
(3) Biología
(2) Química
(1) Física

TABLA 3
Tabla habitual de la reducción.

Medawar sugiere que la auténtica relación de las materias superiores con las inferiores de esta Tabla no es simplemente una relación de reductibilidad lógica, sino más bien comparable a la relación entre las materias que se mencionan en la Tabla 4:

(4) Geometría métrica (euclidiana)
(3) Geometría afín
(2) Geometría proyectiva
(1) Topología

TABLA 4
Diversas geometrías.

La relación fundamental entre las disciplinas geométricas superiores y las inferiores, dadas en la Tabla 4, no es muy fácil de describir, pero no es, ciertamente, una de reductibilidad. Por ejemplo, la geometría métrica, especialmente en la forma de geometría euclidiana, es sólo muy parcialmente reducible a geometría proyectiva, incluso aunque los resultados de la geometría proyectiva son todos válidos en una geometría métrica expresada en un lenguaje con la riqueza suficiente como para utilizar los

[2] Véase *El yo y su cerebro [S. & I. B.]*. y también P. B. Medawar, *Introduction and Intuition in Scientific Thought*. Londres, 1969, pp. 15-19, y «A Geometric Model of Reduction and Emergence», en Ayala y Dobzhansky, *op. cit.*, páginas 57-63.

conceptos de la geometría proyectiva. Así pues, podemos considerar que la geometría métrica es un *enriquecimiento* de la geometría proyectiva. Relaciones similares son válidas para los otros niveles de la Tabla 4. El enriquecimiento es, en parte, un enriquecimiento de conceptos, pero principalmente de teoremas.

Medawar propone que las relaciones entre los niveles consecutivos de la Tabla 3 pueden ser análogas a las que hay entre los de la Tabla 4. Así pues, la química puede ser considerada como un enriquecimiento de la física; lo que explica por qué es parcialmente, aunque no totalmente, reducible a la física; y lo mismo ocurre con los niveles superiores de la Tabla 3.

Las disciplinas de la Tabla 4 son claramente *irreducibles* a las de los niveles inferiores, a pesar de que los niveles inferiores permanecen, en un sentido muy claro, válidos dentro de los niveles superiores y a pesar de que están, de alguna manera, contenidos en los niveles superiores.

Además *algunas* de las proposiciones en los niveles superiores son reducibles a los inferiores.

Considero que las observaciones de Medawar son enormemente sugerentes. Sólo son aceptables, naturalmente, si renunciamos a la idea de que nuestro universo físico es determinista, de que la teoría física, junto con las condiciones iniciales que prevalecen en un momento dado, determinan *completamente* el estado del universo físico en cualquier otro momento. Si aceptáramos este determinismo laplaciano, la Tabla 3 no podría considerarse análoga a la Tabla 4.

Tal como son las cosas, puede considerarse que los niveles superiores de esas dos tablas contienen nuevas y fundamentales hipótesis (nuevos axiomas), no derivables de las hipótesis (axiomas) de los niveles inferiores y nuevos y fundamentales conceptos, no definibles en términos de los conceptos de los niveles inferiores.

Contrariamente a esto, la idea del reduccionismo es que nada intrínsecamente nuevo entra en los niveles superiores.

Así, si formalizamos (axiomatizamos) nuestras hipótesis físicas, entonces, según el reduccionismo, todo concepto aparentemente nuevo debería ser reducible (definible) en términos de los conceptos de la física, y, por tanto, evitable en principio; y toda hipótesis aparentemente nueva debería ser, en presencia de esas definiciones, lógicamente deducibles de las hipótesis básicas del sistema de la física, formalizado y axiomatizado.

IV

Ahora bien, existen razones lógicas para dudar de que este programa reduccionista, que puede describirse en términos puramente lógicos, pueda realizarse ni siquiera en principio. Mencionaré algunas de esas razones.

Consideramos un programa similar, el programa de reducir las matemáticas a lógicas: el programa que culminó en los *Principia Mathematica* de Whitehead y Russell, un logro espléndido, pero, a juicio de los matemáticos más competentes, también un fracaso, al menos en lo que se refiere a los aspectos reduccionistas del programa. La lógica pura juega, en efecto, un papel de una importancia inmensa en las matemáticas. Pero las matemáticas son más ricas que la lógica (funcional). Esto puede verse en los descubrimientos de Gödel: en todo sistema axiomático de teoría de números surgen problemas que no pueden ser decididos lógicamente en ese sistema axiomático, sino sólo en uno más fuerte. En este sistema más fuerte surgen problemas nuevos, pero exactamente análogos. Así pues, necesitamos una serie infinita de sistemas axiomáticos crecientes, e incluso la reducción de uno de esos sistemas axiomáticos (incompletos) de teoría de números, a lógica, no sería una reducción completa en el sentido del programa reduccionista.

Existe también el problema de las definiciones. El punto clave de una definición formal en el sentido de un programa reduccionista es que sirve como *mera abreviatura*. W. V. Quine, por ejemplo, después de introducir algunas definiciones en su sistema de lógica matemática, las cuenta como sigue:

> Tales convenciones de abreviaturas reciben el nombre de definiciones formales... Definir formalmente un signo es adoptarlo como expresión taquigráfica de algún tipo de anotación que ya se tiene a mano... Definir un signo es mostrar cómo evitarlo [3].

Este es el tipo de definición que el reduccionista tiene en mente; porque quiere mostrar que no hay elemento intrínsecamente nuevo, no hay elemento irreductible que entre en los niveles superiores: *todo* es reducible al nivel inferior, la física; aunque, debido a las situaciones físicas complicadas, o constelaciones, las defi-

[3] W. V. Quine, *Mathematical Logic*, edición revisada, 1951, p. 47.

niciones que abrevian se hacen necesarias (en razón de lo que Mach llamó la «economía del pensamiento»).

Llamemos a esas definiciones que sólo abrevian «definiciones no creativas». Porque existen también otras definiciones, las definiciones creativas. Formalmente no se distinguen de las no creativas, pero cumplen una función enteramente diferente —la función de los axiomas, o de las nuevas hipótesis—, y su uso no es permisible, por tanto, en una tentativa de reducción.

V

Las definiciones creativas y no creativas pueden describirse de la siguiente forma:

Sea *S* el nuevo símbolo que cierta definición formal introduce. Si la definición es no creativa, es meramente abreviativa, entonces todos los *nuevos teoremas* —es decir, aquellos teoremas cuya derivación permite la definición y que no pueden ser derivados sin la definición— contendrán el símbolo *S*; y la definición nos permitirá eliminar el símbolo *S* de cada uno de esos teoremas nuevos. Sin embargo, si la definición es «creativa», entonces habrá teoremas que no contengan el símbolo *S* pero que no serán derivables de los axiomas en ausencia de esa definición que introduce *S*.

Las primeras impresiones pueden sugerir que no pueden existir tales definiciones creativas. Sin embargo, pueden y existen; y algunos de los hechos que las conciernen afectan fuertemente al programa reduccionista.

En 1963 publiqué un estudio, «Creative and non Creative Definitions in the Calculus of Probability» [4]. Tomé el cálculo de la probabilidad como objeto de mi estudio por varias razones, principalmente porque me ofrecía un sistema axiomático que pensé que conocía bastante bien, y porque estaba bastante familiarizado con los métodos para probar si un axioma (o una definición) era capaz o no de dar lugar a nuevos teoremas; es decir, teoremas no derivables del resto del sistema axiomático.

Los principales resultados del estudio que nos interesan aquí son los dos siguientes (de los que se dan ejemplos en ese estudio):

Si introducimos una definición puramente abreviativa, o no

[4] *Synthese*, 15, 1963, pp. 167-186.

creativa, en un sistema axiomático, entonces *esta definición puede convertirse en una definición creativa:*

a) omitiendo uno de los axiomas,
b) añadiendo un axioma nuevo.

Así pues, no podemos tener nunca la certeza de si una definición es creativa o no es creativa a menos que nuestro sistema axiomático esté estrictamente fijado.

Ahora bien, esto es, naturalmente, de la mayor importancia desde el punto de vista del «fisicalismo»; es decir, de la tesis de que (por lo menos) la totalidad de la química y de la biología pueden reducirse a física: no podemos decir de ninguna definición que, en apariencia, sea puramente abreviativa, que lo es realmente, a menos que el sistema axiomático de la física con el que operamos (al que queremos reducir los sistemas superiores) esté formalizado y fijado con precisión.

Pero aun entonces el carácter de alguna definición puede quedar en duda. No existe un método rutinario para decidir si una definición dada es creativa o no en relación con un sistema axiomático dado.

Me parece que esto muestra que, desde un punto de vista puramente lógico, el programa del reduccionismo es realmente muy vago. Naturalmente, desde el punto de vista de entender intuitivamente una ciencia —un aspecto muy importante, algo vago— incluso las reducciones parciales pueden seguir siendo deseables e importantes.

VI

Todo esto atañe a la historia del materialismo y a la historia de la auto-transcendencia del materialismo.

La principal idea tras el materialismo, el programa de investigación materialista, como yo lo llamo [5], era tratar de explicar todas las cosas —el Todo, el universo ordenado, el Cosmos— en términos de las familiares y bien conocidas propiedades de la materia.

---

[5] Véase el «Epílogo Metafísico» a *Teoría cuántica y el cisma en Física [Q. T. & S., Ph.]*, volumen III del *Post Scriptum* a la *Lógica de la investigación científica.*

Había, en general, dos programas materialistas de investigación. Uno, que se remontaba a Parménides, consideraba el mundo lleno, lleno de materia; y llevó a la mecánica del continuo. El otro, tomando «los átomos y el vacío» como consigna, consideraba el mundo vacío en gran medida. Los dos programas llevaban a una visión del mundo como una inmensa máquina mecánica: bien fuera de vértices o de átomos. Pero era esencial a los dos programas que el mundo se expresara en términos de las propiedades conocidas de la materia.

Esta exigencia esencial es, naturalmente, reduccionista. En este aspecto, el materialismo y el reduccionismo son uno y el mismo programa. Fue un programa enormemente importante y fructífero, uno que realmente se convirtió en la ciencia de la naturaleza. Sin embargo, se transcendió a sí mismo, gracias a la tradición crítica de la ciencia, que resultó más fuerte que la tradición ideológica.

Así pues, en lugar de aquellas propiedades familiares que debían haber explicado las cosas, como se esperaba en el programa original, ahora tenemos leyes abstractas y nada familiares: y la conducta familiar de la materia se explica por medio de las nada familiares fórmulas matemáticas abstractas. Por ejemplo, la idea sumamente satisfactoria, intuitivamente, de la conservación de la materia ha sido sustituida por una ley sumamente abstracta de la conservación de la energía; y a la propia materia se la considera sólo como una fórmula de esa energía abstracta.

Pero este proceso del materialismo no transcendente comenzó anteriormente con Newton y las fuerzas newtonianas, con Faraday y Maxwell, y con Einstein y la idea de los campos. Y con ideas como la de una probabilidad intrínseca de desintegración atómica (vida media) [6].

VII

Ninguno de esos esfuerzos reduccionistas explican la creatividad del universo: la vida y sus increíbles complejidades y riqueza de formas. De hecho, antes de Darwin un reduccionista no podía hacer más que cerrar los ojos ante el problema del designio en la naturaleza. Después de la publicación, en 1859, de *El origen de*

[6] Véase *El yo y su cerebro [S. & I. B.]*, especialmente capítulos P1, P3 y P5.

las especies, hubo un argumento —la selección natural—, un argumento realmente muy fuerte, que podría ser usado por el reduccionismo. Los reduccionistas no necesitaban ya cerrar los ojos para no ver el problema del designio; por el contrario, ahora podían utilizar el problema del designio en favor del reduccionismo.

El programa de reducción darwiniana recibió su mayor aliento con el éxito de Watson Crick. No es extraño que la biología molecular se convirtiera no sólo en una rama de la ciencia de un crecimiento extremadamente rápido, sino, al mismo tiempo, casi en una ideología.

VIII

Quiero decir unas palabras sobre otro acontecimiento recientemente muy estimulante que tiene gran importancia para la cuestión de la evolución de la vida: el desarrollo de la termodinámica de sistemas abiertos distantes del equilibrio.

«Termodinámica» quiere decir el flujo del calor y las fuerzas responsables de él. El calor fluye, como todo el mundo sabe, desde el cuerpo o región más caliente al más frío y este movimiento tiende al equilibrio, en que el flujo cesa. La termodinámica como ciencia trata de describir esto; y la mecánica molecular correspondiente, llamada mecánica estadística, proporciona una explicación materialista y reduccionista de éxito.

Las dos primeras leyes de la termodinámica son la ley de la conservación de la energía y la ley que afirma que la entropía sólo puede aumentar. Expresada en los términos de la interpretación de Boltzmann de la entropía como desorden molecular, la segunda ley dice que el desorden molecular de un sistema cerrado sólo puede aumentar hasta que alcanza su máximo: el desorden total.

Esta ley del incremento del desorden, interpretada como principio cósmico, hizo la evolución de la vida incomprensible, aparentemente incluso paradójica. Porque la evolución de la vida muestra tendencia al alejamiento del desorden boltzmanniano.

Durante mucho tiempo se ha sospechado que la solución a esta aparente paradoja está relacionada con el hecho de que todo sistema viviente, e incluso la tierra entera, con su flora y fauna creciendo y desarrollándose, es *un sistema abierto*.

Naturalmente, la segunda ley (y la interpretación que dio

Boltzmann de ella) no es aplicable a los sistemas abiertos; así que ahí parecía haber una posibilidad de progresar.

Ahora ya se han hecho progresos espectaculares. No puedo contar aquí cómo, pero quiero mencionar los resultados más importantes, debidos principalmente a Prigogine [7]. Son, brevemente, que los *sistemas abiertos en un estado distante del equilibrio* no muestran tendencia a incrementar el desorden, incluso aunque produzcan entropía. Pero pueden exportar esta entropía a su entorno, y pueden incrementar, en vez de disminuir, su orden interno. Pueden desarrollar propiedades estructurales y, por tanto, hacer lo contrario, quedar en un estado de equilibrio en el que ya no puede ocurrirles nada estimulante.

Quizá el ejemplo más sencillo sea el de una tetera hirviendo en un hornillo. Es un sistema abierto, en el sentido de que fluye a ella mucha energía desde el fondo y fluye hacia fuera de ella por la parte de arriba y por los lados.

Dentro del sistema, aparecen fuertes diferencias de temperatura, al contrario de lo que ocurriría en un sistema cerrado. Producen no sólo un flujo de calor, sino también repetidas corrientes de agua, y, cuando el agua empieza a hervir, tenemos incluso la producción de estructuras materiales visibles de un tamaño bastante característico: burbujas de vapor. Esas burbujas de vapor distan mucho de ser iguales, pero hay una especie de tamaño medio: un efecto probabilista o estadístico (tales propensiones dependen de la situación total: temperatura del hornillo, tamaño y forma de la tetera, flujo de calor...). Además, hay una división del agua en dos estados —agua líquida y vapor—, y es claramente un problema probabilista si un grupo de moléculas adoptará un estado o el otro durante la unidad de tiempo siguiente: nos enfrentamos aquí (igual que en toda la termodinámica) con efectos probabilistas, con una parte de la física que no es determinista.

Prigogine está desarrollando esta parte de la física, tanto teórica como experimentalmente, y ahora es claro que los sistemas abiertos en un estado distante del equilibrio pueden construir nuevas estructuras en lugar de moverse hacia un estado de equilibrio, de maximización de la entropía, de desaparición de la estructura: de esa muerte del calor que desde hacía tanto tiempo se predecía para el universo.

---

[7] Ilya Prigogine, *From Being to Becoming: Time and Complexity in the Phisical Sciences* (San Francisco, 1980), especialmente pp. 88-89.

IX

La obra de Prigogine puede considerarse una estimulante reducción fisicalista, al menos en el sentido de que da los primeros pasos hacia una comprensión física de la evolución de las estructuras superiores (que parece ser un aspecto bastante obvio de la evolución de la vida sobre la tierra). Puede así abrir paso a la comprensión de la razón por la que la creatividad de la vida no *contradice* las leyes de la física.

Pero aunque éste es un paso en dirección reduccionista, está infinitamente alejado de una reducción de las propiedades creativas de la vida.

Miremos o no al universo como a una máquina física, tenemos que enfrentarnos al hecho de que ha producido vida y hombres creativos, que está abierto a sus pensamientos creativos y que ha sido cambiado físicamente por ellos. No debemos cerrar los ojos ante este hecho ni permitir que nuestra apreciación de los éxitos fraguados por los programas reduccionistas nos impida ver que el universo que alberga la vida es creativo en el mejor sentido: creativo en el sentido en el que lo han sido los grandes poetas, los grandes artistas, los grandes músicos, y también los grandes matemáticos, los grandes científicos y los grandes inventores.

Sir KARL (RAIMUND) POPPER nació en Viena en 1902. Estudió matemáticas, física, psicología, educación, historia de la música y filosofía en la Universidad de Viena entre 1918 y 1928. Al mismo tiempo trabajó como aprendiz de ebanista y como maestro. En 1978, cincuenta años después de que obtuviera su doctorado, la Universidad de Viena «renovó» su título en una solemne ceremonia y le hizo doctor *honoris causa* por ciencias naturales.

En 1934, siendo todavía maestro de escuela, publicó *Logik der Forschung*, un libro que se convirtió en un clásico después de su traducción al inglés (*La lógica del descubrimiento científico* se publicó en 1959). Ahora está traducido a muchos idiomas y, después de cuarenta y siete años, sigue reimprimiéndose con frecuencia.

Popper, que vive en Inglaterra, ha dado conferencias en Europa, Nueva Zelanda, Australia, India, Japón y, desde 1950, en que dio las «William James Lectures in Philosophy» en la Universidad de Harvard, las ha dado a menudo en América. Entre sus libros están *La sociedad abierta y sus enemigos* (por el que recibió el Lippincott Award of the American Political Science Association), *El desarrollo del conocimiento científico*, *La miseria del historicismo*, *Conocimiento objetivo* y *Búsqueda sin término*. Junto con Sir John Eccles publicó *El yo y su cerebro*.

Posee títulos honorarios de las Universidades de Chicago, Denver, Warwick, Canterbury (Nueva Zelanda), Salford, The City University (Londres), Viena, Mannheim, Guelph (Canadá), Francfort, Salzburgo, Cambridge y Oxford.

Es Fellow de la Royal Society (Londres) y de la British Academy; Foreign Honorary Member, American Academy of Arts and Sciences; Membre de l'Institut de France; Socio Straniero dell'Accademia Nazionale dei Lincei; Membre de l'Académie Internationale de Philosophie des Sciences; Associate, Académie Royale de Belgique; Membre de l'Académie Européenne des Sciences, des Arts et des Lettres; Honorary Member, the Royal Society of New Zealand; Membre d'Honneur, Académie Internationale d'Histoire des Sciences; Ehrenmitglied, Deutsche Akademie für Sprache und Dichtung.

Es también miembro honorario del Harvard Chapter of Phi

Beta Kappa; Ehrenmitglied, Allgemeine Gessellschaft für Philosophie in Deutschland; Honorary Fellow, London School of Economics and Political Science, y Honorary Fellow, Darwin College, Cambridge.

Ha recibido el Prize of the City of Vienna for Moral and Mental Sciences, the Sonning Prize of the University of Copenhagen, el Dr Karl Renner Prize of the City of Vienna y el Dr Leopold Lucas Prize of the University of Tübingen.

Ha recibido la Grand Decoration of Honour in Gold (Austria); le Gold Medal for Destinguished Service to Science of the American Museum of Natural History, Nueva York; la Ehrenzeichen für Wissenschaft und Kunst (Austria), y la Ordre Pour le Mérite (República Federal de Alemania).

En 1965 recibió el título británico de caballero de la Reina.

WILLIAM WARREN BARTLEY III, graduado por Harvard y por la Universidad de Londres, es ex-alumno, colega y, desde hace mucho tiempo, asociado de Sir Karl Popper. Ha sido profesor de Lógica en la London School of Economics, profesor de Historia de la Filosofía de la Ciencia en el Warburg Institute y S. A. Cook Bye-Fellow en Gonville and Caius College, Cambridge. Anteriormente fue catedrático de Filosofía y de Historia de la Filosofía de la Ciencia en la Universidad de Pittsburgh, y, desde 1970, es catedrático de Filosofía en la Universidad del Estado de California, en Hayward, donde, en 1979, recibió el título de Profesor Distinguido en el Sistema de Diecinueve Campus de la Universidad del Estado de California.

INDICES

Compilados por Stephen Kresge y Nancy Artis Sadoyama

Las referencias en cursiva indican los pasajes de especial importancia. El número de páginas seguido de *t* indica que el término se discute en el lugar referido; *n* significa «nota», y *c* (en el Indice de Nombres) significa «citado». Se aconseja al lector que consulte los índices de los demás tomos del *Post Scriptum*.

INDICE DE NOMBRES

Adams, W., 20.
Agassi, J., 19.
Agustín, San, 29, 105 y *n.*
Ampère, A. M., 159.
Anderson, C. D., 161.
Aston, F. W., 184.
Ayala, F. J., 152 *n,* 186 *n.*

Bartley III, W. W., 20, 21.
Beethoven, L. van, 148-149.
Berkeley, G., 137, 175.
Bohr, N., 160-163, 183-185.
Boltzmann, L., 192-193.
Bolzano, B., 155.
Bondi, H., 164.
Bonhoeffer, K. F., 40 *n.*
Born, H., 26 *n.*
Born, M., 26 *n,* 123 *n,* 160.
Brunswick, E., 177.
Buffon, G., 171.
Bühler, K., 106 *n,* 143 *n,* 174 y *n.*
Bunge, M., 160 n.

Calvino, J., 29.
Cantor, G., 155.
Carnap, R., 119.
Cassirer, E., 44 *n.*
Cotes, R., 157.
Crick, F., 184, 192.
Chadwick, J., 160, 161 *n.*

Dahrendorf, R., 20.
Darwin, C., 170, 191-192.
De Broglie, L., 160.
Descartes, R., 49 *n,* 105 y *n,* 108, 156-157, 159, 160-162, 173, 175, 179.
Diderot, D., 171.
Diels, H., 70 *n.*
Dirac, P. A. M., 160, 164.
Dobzhansky, T., 152 *n,* 186 *n.*
Duhem, P., 36 *n c,* 62 *n.*

Eccles, J. C., 40 *n,* 135 *n,* 168, 172, 173 y *n.*
Eddington, A. S., 160, 161 y *n.*
Einstein, A., 23, 26 y *n,* 27 *n,* 28 *n,* 53, 55, 61 y *n,* 62, 80 *n,* 81 y *n,* 82 *n,* 112 y *n,* 113 y *n,* 114-115, 118, 129 y *n,* 145-146, 155, 158-159, 160 y *n* y *c,* 166 *n,* 191.

Eötvös, R. von, 158.
Euclides, 141, 155, 180, 186.
Euler, L., 45.

Faraday, M., 159, 166 *n*, 191.
Freed, J., 19.
Freed, L., 19.

Galileo, 153, 166.
Gallie, W. B., 103 *n*.
Gamow, G. A., 164.
Geulincx, A., 175.
Gödel, K., 27 *n*, 64 *n*, 81 *n*, 82 *n*, 149, 182, 188.
Goldstine, H. H., 56 *n*, 66 *n* y *c*.

Haas, A., 184 y *n*.
Hadamard, J., 62 y *n*, 63 y *n c*, 73.
Haeckel, E., 171.
Haldane, J. B. S., 70 *n c*, 105 y *c* y *n*, 106, 108-109, 110-112.
Hayek, F. A. von, 144-145 y *n*.
Heisenberg, W., 59 y *n*, 76, 79, 160.
Heitler, W., 161, 185.
Helmholtz, H., 159.
Hertz, H., 118, 159.
Hobbes, T., 27, 43 y *c* y *n*, 46, 49.
Homero, 148.
Hume, D., 21-22, 27, 43 y *c* y *n*, 46, 49. 66, 143.
Huygens, C., 80 *n*.

Jeffreys, H., 119.
Jenófanes, 70.
Johnson, W. E., 137.
Jordan, P., 129, 164.

Kant, I., 23 *n*, 27, 31, 44 y *c* y *n*, 46, 49, 51, 61, 62 *n*, 70, 71 y *n*, 72, 73 y *c* y *n*, 103.
Kepler, J., 30, 153, 166.
Keynes, J. M., 119.
Khinchine, A. I., 119 *n*.
Körner, S., 103 *n*.
Krantz, W., 70 *n*.
Kuratowski, K., 154.

Lakatos, I., 119 *n*.
Lanczos, C., 158.
Landé, A., 122, 123 y *c* y *n*, 124 y *c*, 125-129, 137.
Laplace, P., 22 y *c* y *n*, 23, 28 *n*, 53, 54 *n*, 55, 57-58, 61, 62 *n*, 64, 70, 84-85, 91, 94, 101, 144-148, 151, 187.
Leibniz, G. W., 155, 161, 171, 175.
LeSage, G. L., 156.
Lewis, H. D., 118 *n*.
London, F., 161, 185.
Lorentz, H. A., 166 *n*.
Lutero, M., 29.

Mach, E., 158, 175, 189.
Malebranche, N., 175.
Maxwell, J. C., 23, 53, 55-56, 122 *n*, 145, 154 *n*, 159, 166 *n*, 191.
Mayr, E., 170.
Medawar, P. B., 186 *n*, 187.
Meyerson, E., 153 y *n*.
Mill, J. S., 27.
Miller, D., 20, 169.
Millikan, R. A., 160 y *c* y *n*.
Minkowski, H., 81.
Monod, J., 167 y *c* y *n*, 168 y *c*, 169, 172.
Mozart, W. A., 45, 51, 65.
Musgrave, A. E., 20.

Narlikar, V. V., 77 *n*.
Newton, I., 23, 30-31, 49, 53-56, 70, 71 y *n*, 72-73, 75, 76, 77 *n*, 79, 118, 144-145, 146, 153-159, 161, 162, 166 y *n*, 191.

Olsen, R. E., 174 *n*.

Parménides, 26 *n*, 113, 191.
Paul, A. M., 174 *n*.
Pauli, W., 26 *n* y *c*, 112 *n*, 126 *n*, 160-161.
Pauling, L., 163.
Pecker, J. C., 164 *n*.
Peirce, C. S., 22.
Pendse, C. G., 77 *n*.
Petersen, A. F., 20.
Platón, 148, 154.
Poincaré, H., 86 *c*.
Priestley, J., 43 y *n*.
Prigogine, I., 192, 193 y *n*, 194.

Quine, W. V., 188 y *c* y *n*.

Roberts, A. P., 164 *n.*
Ross, A., 174 y *n.*
Russell, B., 20, 155, 188.
Rutherford, E., 159.

Salam, A., 162 *n.*
Schilpp, P. A., 27 *n, 81 n,* 82 *n.*
Schlick, M., 23 *n,* 27 y *c,* 28 *c* y *n,* 49, 65, 66 *n.*
Schopenhauer, A., 27.
Schrödinger, E., 27 *n,* 146 *n,* 160.
Schubert, F., 136.
Sellars, W., 106 *n.*
Settle, T., 20.
Shearmur, J., 20.
Soddy, F., 184.
Sperry, R. W., 173.
Spinoza, B., 27, 43 y *n,* 46, 175.

Tarski, A., 91 *n.*
Thomson, J. J., 159, 184.
Turing, A. M., 96 *n.*

Urbach, F., 184.
Urey, H. C., 184.

Vigier, J. P., 164 *n.*
Von Neumann, J., 56 *n,* 66 *c* y *n.*

Watkins, J. W. N., 20.
Watson, J., 184, 192.
Weber, W., 159.
Weinberg, S., 162 *n.*
Weyl, H., 83 *n.*
Whitehead, A. N., 188.
Whorf, B. L., 178.
Wiener, P. P., 154.
Wittgenstein, L., 66.
Woodger, J. H., 114.

Yukawa, H., 161 y *n.*

INDICE DE MATERIAS

Acción, libertad de, *véase* Libertad humana.
Agua pesada, descubrimiento de, 185.
Aleatoriedad, hipótesis general de, 124.
Aprendizaje, 37-38, 44.
Aproximaciones, en ciencia, 74-75, 84-85, 92-93;
  series infinitas de, 68-99;
  y modelos, *sección 17.*
Armonía, preestablecida, de la teoría de las tablas periódicas, 165 y *n;* segundo tipo, 165;
  y panpsiquismo o hilozoísmo, 171.
  Véase también Reducción de la química a física.
Atomismo, 190-191.
Auto-predicción, imposibilidad de, *sección 22;*
  y prueba de que la calculadora no puede predecir resultados de sus propios cálculos, 92-101;
  y tarea de predicción, 95-101.
  Véase también Conocimiento, aumento de; Predictibilidad.

Biología, 26 *n,* 38-39, 152, 191-192.

Calculadora, argumento de la máquina, 129-132, 146, *sección 22.*
Cambio, experiencia del, 26 *n*-27 *n;* realidad del, 27 *n,* 113-115.
Caos, molecular, principio de, *véase* Aleatoriedad.
Capacidad de dar razón, 35;
  principio de, *secciones 3 y 4;*
  problema de, y física clásica, 37-38, *sección 17;*
  sentido más fuerte de, y mediciones, 36-37, *sección 17;*
  sentido más hábil de, *sección 17;*
  y argumento desde el estudio de la conducta, 38-39, 41;
  y argumentos desde la psicología, 43-48;

y auto-predicción, 104;
y determinismo «científico», 37.
y el arte de la ultra-simplificación sistemática, 68;
y la dinámica de Newton, 74-79;
y la precisión de las condiciones iniciales, 35-38;
y las tareas de predicción, *secciones 3 y 7;*
y necesidad de un modelo, 68;
y temperatura crítica, 40;
y transmisión nerviosa, 40-41.
Capacidad de «dar y recibir patadas» y realidad, 137, 140.
Capacidad de predicción, perfeccionamiento de, y determinismo «científico», 38-39.
Carga de la prueba, 50, *secciones 9 y 25.*
Cartesiano, programa de reducción, 156-157, 159.
Causalidad, idea de, popular o según el sentido común, *sección 1,* 45-47;
ley de la causación universal, 33;
principio de, 28.
Cerebro, su emergencia, 143;
y sensibilidad a los efectos-masa, 40-41.
Ciencia, objetivos de, 69-71;
teoría de la, como faro, 69 *n.*
Código genético, 168-169.
Completud, 62, 149-150.
Comportamiento, *véase* Conducta.
Computadoras, *véase* Calculadora, argumento de la máquina.
Conciencia, 26 *n,* 48;
emergencia de, 143, 173.
Conducta, argumento desde el estudio de, *sección 4;*
y capacidad de dar razón, 38-39, 43-44.
Conductismo filosófico (o teoría de la identidad), 138;
y negación del Mundo 2, 175-176.
Conocimiento, aumento de, y predicción, 41, 44-46, *secciones, 7, 15, 20 y 25;*
cómo su existencia crea indeterminismo, *secciones 23, 24 y 31;*
formulación científica y Mundo 3, 136;
incompletud de, y cambio, 146;
insuficiencia de, y probabilidad, 126-127;
objetivo, 139-140, 150-151;
y aceptación de teoría no aceptada anteriormente, *sección 21.*
Conocimiento previo, 54-55, 110.
Conservación, leyes de, 191.
Contenido del pensamiento, 139-140; *véase también* Mundo 3.
Contrastabilidad, de las teorías probabilistas, 120-122, *secciones 29 y 30;*
de las propensiones, 118;
y debilidad lógica, 31-32;
y preferencia por teorías contrastables, 67-68.
Cosmología, física, 21.
Creatividad, humana, 23-24, 44-46, 134, 144-145, 148, 152, 194.
Crítica, exposición o discusión, 31-32, 56-57, 107-109, 139.
Cuestiones de-por-qué, *sección 2.*
Cuestiones estadísticas y teorías estadísticas, *secciones 28 y 29.*

Decisiones, causadas o incausadas, 45;
no pueden ser sustituidas por predicciones científicas, 104.
Definiciones, creativas y no creativas, 189 *t*-190.
Designio, en las naturalezas, problema de, 191-192.
Determinismo, determinista, argumentos científicos en favor de, 1;
argumentos filosóficos en favor de, 25;
argumentos lógicos contra, 65-66;
argumentos metafísicos contra, 66;
carácter del mundo, no entrañado por carácter, de teoría, 56;
clásico, o del Mundo 1, *sección 10,* 144;
concepción del mundo, *sección 8;*
distinto de causalidad, 28, 45-46;
doctrinas metafísicas de, *sección 25;*
hace superfluo el futuro, 114;
idea de, según el sentido común, 25, 32-34;
posibilidad de argüir sobre, 28;
Schlick sobre, 27-28 y *n.*
y física clásica, *sección 10;*

y la carga de la prueba, *sección 9,* 110-112;
y la cuchilla de Landé, *sección 29;*
y mecánica cuántica, 145-147;
y reduccionismo, 152, 187;
y teoría de las propensiones, *sección 26.*
Véase también Determinismo «científico»; Determiniso metafísico; Determinismo religioso; Indeterminismo.
Determinismo «científico», 22 *t,* 25-26 *t,* 28-29 *t,* 34 *t, sección 1;* Capítulo II, 36-37, 46-47, 48, 66;
apelación al éxito de la ciencia, 57;
argumento desde, y carácter aproximado del conocimiento, 79-80;
argumentos lógicos contra, 66;
argumentos metafísicos contra, 66;
cambio de opinión de Einstein con respecto a, 26 *n* y 27 *n, sección 26;*
dos definiciones de, *sección 12;*
e imposibilidad de la autopredicción, 100-102;
firme creencia de Kant en, 31, 44, 49;
laplaciano, 22-23, 144-148;
no entrañado por la mecánica de Newton, 30-31;
refutación de, *secciones 17 y 23;*
refutado por el argumento de Haldane, Descartes y San Agustín, *sección 24;*
refutado por el resultado de Hadamard, *sección 14;*
requiere capacidad de dar razón, *secciones 3 y 7;*
supuesto carácter científico de, 22;
surge de la crítica de la analogía de relojes y nubes, 32;
tercera versión, donde descripción después del suceso entraña lógicamente predicción, 103-104;
tercera versión y capacidad de dar razón, 103-104;
tiene que ignorar las diferencias entre las funciones superiores e inferiores del lenguaje, 108;
versión más fuerte de, 36, 60-61, 62;
y cuatro funciones del lenguaje, 106-109;
y predictibilidad desde dentro, *sección 11;*
y relatividad especial, 85-86;
y teorías deterministas *prima facie,* 56, *sección 13.*
Véase también Determinismo; Determinismo metafísico; Indeterminismo.
Determinismo metafísico, 31-32 *t,* 41, *sección 25;*
entrañado por el determinismo «científico» y el religioso, 32;
incontrastabilidad de, 31-32;
los argumentos más fuertes en favor de, provienen del determinismo «científico», 32, 50;
posibilidad de argüir sobre, *sección 25;*
rechazo de, *sección 26;*
y el argumento de Haldane, 111.
Véase también Determinismo; Determinismo «científico»; Indeterminismo.
Determinismo religioso, 28-31, *secciones 12 y 25.*
Dios, 29, 110;
cuestión de la omnisciencia de, 29, 30;
cuestiones de la omnipotencia de, 29 y *n.*
DNA, descubrimiento de la estructura de, 184.
Doppler, efecto, 76-77.
Dualismo, 139-140. *Véase* Mundos 1, 2 y 3.

Efectos-masa, temperatura como, y capacidad de dar razón, 40-41.
Einstein, teoría de la relatividad especial de, destruye la identidad esencialista de la masa inerte y la masa gravitatoria, 158;
y determinismo «científico», 85.
Véase también Relatividad, teoría especial de.
Einstein como Parménides, y determinismo metafísico, 26 *n*-27 *n, sección 26.*

Einstein sobre determinismo, 26 y *n.*, 61, 112 *n*-113, 112-115;
sobre las partículas elementales, consideradas como condensaciones del campo electromagnético, 160 y *n.*
Emergencia, y propiedades emergentes, 166 y *n.*
Entropía, en sistemas abiertos y cerrados, 192-193.
Enunciados de probabilidad, de sucesos únicos, *sección 29;*
e hipótesis general de aleatoriedad, 125;
y la cuchilla de Landé, 123, 124-127.
Epifenomenalismo, 175.
Esencialismo, *véase* Reduccionismo filosófico.
Estabilidad, estadística, 121-122;
problema de, 55, *secciones 12 y 14.*
«Estado» del sistema newtoniano, 53;
cuestión de si un sistema estará alguna vez en un determinado estado, *sección 12.*
Evolución, teoría de, en física y en biología, 170-172;
y termodinámica, 191-194.
Exactitud, *véase* Precisión.
Explosión, capacidad de dar razón y el principio de todo-o-nada, 40.

Falibilidad, 150.
Filosofía moderna, y análisis verbales, 22-24, 65-66, 134, 152, 153 y *n.*
Física, clásica, argumentos en favor del determinismo en, *secciones 8 y 10;*
cuestión de su capacidad de dar razón, 37, *sección 17;*
determinismo *prima facie* de, 22, 26, *sección 10;*
diferencias con la física cuántica, 53;
es indeterminista, 146;
similar a la física cuántica respecto a la capacidad de dar razón, 78-99.
Fisicalismo, *véase* Conductismo filosófico.
Fisiología, procesos de, 44-45;
supuestamente determinista, 39.
Flecha del tiempo, 80 *n.*
Véase también Pasado y futuro.
Fluctuaciones, 117.
Futuro, apertura de, 83;
cognoscibilidad de, 29 *n*, 71-73;
contenido o no en el pasado y en el presente 27 *n*, 54, 72;
predictibilidad de, 29;
se hace superfluo en el determinismo, 114;
supuesto inmutable, 31.

Gödel, argumento de, contra la realidad del tiempo y del cambio, 27 *n;*
teorema de, 149-50; 182, 188.

Hadamard, resultado de, *sección 14;*
y argumento de Gödel, 64 *n.*
Hafnio, descubrimiento de, 183-184.
Haldane, argumento de, y la refutación del determinismo «científico», *secciones 24 y 25;*
y determinismo metafísico y religioso, 110.
Heisenberg, argumento de, contra el determinismo, 59 y *n;*
fórmulas de indeterminación, 179.
Historia, humana, curso futuro de, impredictible, 87-88.
Historicismo, doctrina de, y predicción, 87-88, 91;
refutación de, 87.
Hobbes, concepción determinista de, del mundo físico, 49;
y teoría de Newton como combinación de, 49.
Hombre, naturaleza de, más que predictor, 129-131.

Idealismo, 27 *n*, 79, 114-115, 138.
Identidad, teoría de, *véase* Conductismo filosófico.
Indeterminismo, Capítulo III; comparación con el mundo de los noumenos de Kant, *sección 16;*
doctrina metafísica de, *sección 25;*
espacio para, en física, 23-24;
no basta, 134-135, 147, 151;
razones en favor de, 25, 50-51;
y conocimiento, creado por su existencia en el mundo físico, 104-105, *sección 31;*

y creatividad, 65;
y la carga de la prueba, *véase ibid.;*
y libertad humana, necesario pero insuficiente para, 134-135, 147-148;
y mecánica cuántica, 145-147;
y unicidad del mundo, 70, *sección 16.*
Indeterminismo metafísico, incontrastabilidad de, 32, *sección 25.*
Innovación, *véase* Creatividad.
Interaccionismo psico-físico, 176-178. *Véase también* Mundos 1, 2 y 3.
Interferencia, en teoría cuántica, *véase* Heisenberg, argumento de; y Capacidad de dar razón, 75-80.
Intuición, 23-24, 25.

Kant, noumenos de, comparados con la visión indeterminista de la unicidad del mundo, *sección 16.*

Landé, cuchilla de, *secciones 29 y 30;*
como argumento contra el determinismo, 122-124;
como argumento contra la teoría subjetivista de la probabilidad, 126.
Laplace, demonio de, 53-54 *t,* 57-58, 61, 91, *sección 10;*
requisitos para, 57-59;
y máquina de predecir, 94-95;
y teorías deterministas *prima facie,* 55;
y veredicto de la relatividad especial, 83-86.
Laplace, determinismo de, 22-23, 144-148, 187.
Lenguaje, cuatro funciones de, 106 y *n*-107, 110, 143 y *n*-144, 174 y *n*-175;
teoría del, como faro, 69;
y Mundo 3, 137, 139, 143-144.
Leyes, de la naturaleza, especificación de, 33-34.
Libertad, humana, 21-23, 25, 43-44, 46, 65, 129, 134-135, 145, 151;
de crear, y Mundo 3, 27, 134-135, 145, 147, 148-149;
de evaluar razones, 134-135, 147-148;
y necesidad de apertura causal, 135. *Véase* Mundos 1, 2 y 3.
Libre albedrío, *véase* Libertad humana.
Logros únicos, supuesta predictibilidad científica de, 65-66, 148-149.

Materialismo, monista, *véase* Conductismo filosófico;
programa de investigación materialista, 190-191;
trascendencia de, 191;
Mecánica de Newton, 31, 61-62, 145;
capacidad de dar razón de, 75-79;
como teoría verdadera del mundo, 70;
y reducción científica, 156-159.
Medición, y capacidad de dar razón, 76-79.
Modelos, e infinita retrogresión, *sección 17.*
Motivos, inferencia de, en Hume, 43.
Muchos-cuerpos, problema de, 54-56, 73-74, 75.
Mundos 1, 2 y 3, tres mundos, 135 *t*-136;
e interaccionismo psicológico, 176-178;
interacción entre, 138-139, 149, 175-182;
y emergencia de la conciencia humana, 173-175, 178-181;
y problema cuerpo-mente, 175-178.
Mundo 1 (mundo físico), 135, 147, 148;
confusión de su carácter con el de las teorías del Mundo 3 sobre, 151;
su apertura, 148;
su realidad, 137-138;
supuesta clausura causal hacia Mundos 2 y 3, 145, 147-148, 176.
Mundo 2 (experiencia o pensamiento en el sentido subjetivo), 135 *t;*
realidad de, 137-139, 190-191;
su apertura causal hacia los Mundos 1 y 3, 137-139.
Mundo 3 (pensamiento objetivo, especialmente *productos* de la mente humana), 135 *t*-136;
propiedades autónomas de, 139-142, 151;

realidad de, 139-142;
su apertura causal hacia Mundos 1 y 2, y su apertura intrínseca, 137, 149-151;
y el teorema de Gödel, 149;
y lenguaje, 141;
y números, 140-142.

Naturaleza, leyes de, 29-30;
concepciones determinadas por, y el argumento de Haldane, 105-106.
Noumenos, de Kant, *véase ibid.*
Novedad, intrínseca, y definiciones creativas, 187-190.
Números primos, y Mundo 3, 140-141.

Observador, 83 *n*, 106.
Originalidad, *véase* Creatividad.

Paralelismo psicofísico, 175.
Parménides-Einstein, y determinismo metafísico, 26 *n*-27 *n*, *sección 26.*
Pasado y futuro, 31-32;
argumento de la asimetría entre, *sección 18;*
y la teoría especial de la relatividad de Einstein, *sección 19.*
Película, analogía para el determinismo, 28-29, 56-57;
y Einstein-Parménides, 113-115.
Pluralismo, 137-138.
Potencial de la placa terminal, 40.
Precisión, grado especificado de, 25-26, 30, 34-35, 36, *secciones 3, 4 y 17.*
Predictibilidad, 25, 30, *sección 1;*
auto-predicción, imposibilidad de, 86-88, 88-90, *sección 26;*
de acuerdo con procedimientos racionales, 56;
de logros únicos, 66;
de los relojes y de las nubes, 30, 32;
desde dentro, *secciones 11, 21, 22, 23 y 25,* 59-60, 85-86, 87, 102;
desde fuera, *sección 25;*
grado de precisión de, 35-36, 37;
histórica y el aumento de nuestro conocimiento, *secciones 20 y 21.*
Problema, objetivo, 140-142.
Problema cuerpo-mente, 172-182.
Véase también, Reducción, cuestiones de.
Programas de investigación, biológico, 166 *n;*
materialista, 190-191;
metafísico, 166;
metafísico (cosmológico), 154 y *n;*
pitagórico, 144.
Propensiones, *sección 27;*
como posibilidades físicas, 127-128;
comparadas a fuerzas, 117-118;
fuerzas fluctuantes y, 116-117;
incompatibles con el determinismo, 116;
una teoría física de, *sección 27;*
y conclusiones estadísticas, 129;
y la cuchilla de Landé, *sección 30;*
y situación objetiva, 117.
Pseudo-argumento, 28.
Psicología, argumentos en favor del determinismo desde, *sección 7;*
aceptación de Kant de, 44, 46;
y el principio de poder dar razón, 47-48;
y la idea de Hume de «la inferencia de motivos», 43.
Psicología, de la Gestalt y Mundo 2, 177.

Racionalidad, 108-109, 110-111, 147.
Razón, el poder dar, *véase* Capacidad de dar razón.
Realidad, y Mundo 1, 137.
Realismo, 26 *n*-27 *n.*
Reducción, reduccionismo, cartesiano, 156;
como forma de tentativa científica que obtiene éxito, 153, 158, 183;
como método, 153;
contrastado con explicación, 153-154 y *n*, 155-156;
crítica de Medawar de, 186-187;
cuestiones de, 152-194;
darwiniano, 192;
de biología a química y física, 152, 169-170, 190;
de fracciones racionales a pares ordenados de números naturales, 154-155;
de la conciencia a biología o a química y física, 152, 171-172, 178-180;
de la matemática a la lógica, 188-189;

de la química a la física, 153, 160-167, 170, 184-185;
de los Mundos 2 y 3 al Mundo 1, 180-181;
de mecánica y química a teoría electromagnética de la materia, 160-161;
einsteiniano, 158-159;
filosófico, 153 y *n*-154 y *n*, 167, 181-182, 183;
intentado, de electricidad y magnetismo a mecánica de Newton, 145, 156-157;
lógica de, 186-190;
maxwelliano, 159;
newtoniano, 157-158;
pitagórico, 154;
Prigogine y, 194;
programas de, en física, 155-159;
y cuatro funciones del lenguaje, 173-175.

Relatividad, teoría especial de, no demonio laplaciano en, 84-85;
no determinista *prima facie, sección 19;*
y argumento de Gödel, 27 *n*, 82 *n*-83 *n*.
Véase también Einstein, teoría de; Relatividad, teoría general de.
Relatividad, teoría general de, y argumento de Gödel, 82 *n*-83 *n*.
Relojes y nubes, analogía de, 29-30 y *n*, 32, 50, *secciones 4 y 6;*
argumentos deterministas a partir de, *sección 6;*
y el estudio de la conducta, 38.
Resolución de problemas, y conciencia, 171.
Responsabilidad humana, 22, 134.

Saltos cuánticos, 146 y *n*.
Schlick, su concepción del determinismo, 27-28 y *n*.
Selección natural, *véase* Evolución.
Sentido común, 21, 27 *n;*
idea de causa y «suceso» según el, 21, 34;
idea de determinismo según el, *sección 1.*
Simplicidad, como característica de la teoría científica, 68-69.
Sistema periódico, 183-184.
Sistemas abiertos, 192-194;
universo, 149-151;
y Prigogine, 192-194.
Suceso, sucesos, idea de, según el sentido común, 34;
predictibles e impredictibles, distinción *prima facie, secciones 1, 2 y 6.*
Sucesos fortuitos, 145-146.

Tareas de predicción, y capacidad de dar razón, *secciones 3, 4, 17 y 22.*
Temperaturas críticas, *sección 5;*
y capacidad de dar razón, 41.
Teoría cuántica, 26, 146-147, 159-167, 183-185;
como teoría electromagnética de la materia, 160;
una teoría probabilista de, 53;
y determinismo, 145-146, *sección 10;*
y libertad humana, 146-147;
y reducción de la química a física cuántica, 162-164, 167, 185.
Teoría de conjuntos, axiomática, 155.
Teoría de la probabilidad, una teoría física de, 115-116.
Teoría electromagnética de la materia, 23, 158.
Teorías científicas, argumentos desde las características de las teorías a las características del mundo, 67-68, 69-70;
carácter aproximado de, 79, 86;
como más que instrumentos, 66;
como redes para atrapar el mundo, *sección 15,* 66, 70-71.
como teorías universales, 68-71;
falibilidad de, 67;
implicaciones para el argumento de Haldane de que el mundo es repetitivo, 70 *n.*
Teorías deterministas, *véase* Teorías deterministas *prima facie.*
Teorías deterministas *prima facie,* 22-23, 54-55 *t,* 67-68, *sección 10;*
combinadas con supuestos estadísti-

cos, llevan a una retrogresión infinita, *sección 29;*
Einstein sobre, 26 *n;*
y cuestiones probabilistas, *sección 28;*
y la cuchilla de Landé, 123;
y refutación del determinismo, 102-103;
y teoría de la relatividad especial, 83-86.
Teorías probabilistas, contrastación de, 120-122;
objetivas, en oposición a subjetivas, 119;
y teorías deterministas *prima facie, sección 28.*
Termodinámica, 192 *t.*
Tiempo, flecha del, 80 *n;*
interpretado como coordenada del espacio, 26*n;*
realidad de, 26 *n*-27 *n.*
Todo-o-nada, principio de, *sección 5;* de la transmisión nerviosa, 40.
Transmisión nerviosa, y capacidad de dar razón, 40-41.

Unicidad del mundo, y carácter aproximado de teorías, 70.
Universalidad, 69-70.
Universo, apertura esencial de, 26 *n*-27 *n,* 146, 149-151;
cerrado, parmenídeo, 191;
creatividad de, 191;
teoría de su origen, 163-164.
Universo emergente, 151.

Verdad, como objetivo de la ciencia, 66;
Tarski sobre, 91 *n.*
Vida, origen de, 168, 172-173;
sugerencia de Monod sobre su unicidad, 167-170.
Voluntad, humana, y Hobbes, 43 y *n.*

Yo, conciencia humana del, 172-175, 178-180.

«El universo abierto», uno de los tres volúmenes que integran el tan esperado Post Scriptum que Sir Karl Popper escribió hace más de veinticinco años, contiene el núcleo del argumento de esta obra. En el presente tomo, el autor defiende, en un lenguaje sencillo, la libertad, la creatividad y la racionalidad del género humano. El propio Popper describe este libro como «una especie de prolegómeno a la cuestión de la libertad y la creatividad humanas, cuestión a la que abre espacio tanto física como cosmológicamente de una forma que no depende de análisis verbales».

Además de proyectar una nueva perspectiva cosmológica y física, la obra de Popper exige una nueva comprensión de la historia de la física durante los últimos cincuenta años y de su relación con la física clásica. Después de investigar, y de rechazar, tanto el determinismo «científico» como el metafísico, Popper sostiene que la física clásica, contrariamente a la opinión más común, no presupone determinismo —de la misma manera que no lo presupone la física cuántica—, ni tampoco lo implica. No obstante, Popper considera que el determinismo metafísico continúa subyacente en el pensamiento de muchos teóricos cuánticos contemporáneos, incluidos los adversarios del determinismo.

Sir Karl Popper es doctor «honoris causa» por catorce universidades estadounidenses, británicas, alemanas, austríacas, neozelandesas y canadienses. Es miembro de número u honorario de doce Academias. Algunas de sus obras se han traducido a veintidós idiomas.

La presente edición está preparada por William Warren Bartley III, graduado por Harvard y por la Universidad de Londres. Profesor en las universidades de Londres, Cambridge y Pittsburgh, actualmente es catedrático de Filosofía en la Universidad del Estado de California, en Hayward.

W7-CJF-445

TIT: EL UNIVERSO ABIERTO
AUTOR: POPPER, K.
CODIGO: 1E0431059
PRECIO: 1.950'00